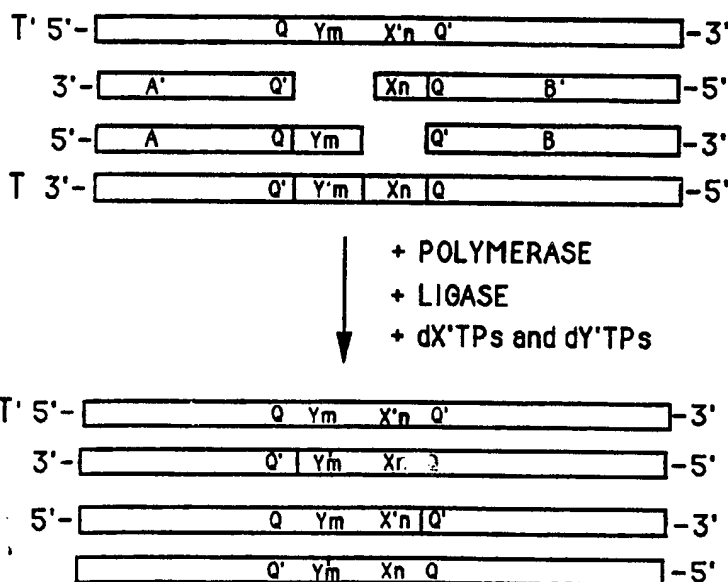




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>5</sup> : C12Q 1/68, C12P 19/34 C07H 15/12, 17/00		A1	(11) International Publication Number: WO 93/00447
			(43) International Publication Date: 7 January 1993 (07.01.93)
(21) International Application Number: PCT/US92/05477		(74) Agents: GORMAN, Edward, Hoover, Jr. et al.; Abbott Laboratories, CHAD 377/AP6D-2, One Abbott Park Road, Abbott Park, IL 60064-3500 (US).	
(22) International Filing Date: 26 June 1992 (26.06.92)			
(30) Priority data: 722,798 28 June 1991 (28.06.91) US		(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE).	
(71) Applicant: ABBOTT LABORATORIES [US/US]; CHAD D377/AP6D-2, One Abbott Park Road, Abbott Park, IL 60064-3500 (US).		Published With international search report.	
(72) Inventors: BIRKENMEYER, Larry, G. ; 5303 W. Leland, Chicago, IL 60630 (US). CARRINO, John, J. ; 5212 Solomon Court, Gurnee, IL 60031 (US). DILLE, Bruce, L. ; 40841 N. Nevelier Drive, Antioch, IL 60002 (US). HU, Hsiang-Yun ; 1763 Cedar Glen Drive, Libertyville, IL 60048 (US). KRATOCHVIL, Jon, David ; 7101 Fifth Avenue, Kenosha, WI 53143 (US). LAFFLER, Thomas, G. ; 1964 Pinehurst Court, Libertyville, IL 60048 (US). MARSHALL, Ronald, L. ; 900 Winthrop Court, Zion, IL 60099 (US). RINEHARDT, Laurie, A. ; 8104 65th Avenue, Kenosha, WI 53142 (US). SOLOMON, Natalie, A. ; 467 Thorndale Drive, Buffalo Grove, IL 60089 (US).			

(54) Title: AMPLIFICATION OF TARGET NUCLEIC ACIDS USING GAP FILLING LIGASE CHAIN REACTION



## (57) Abstract

An improved "gap filling" embodiment of the Ligase Chain Reaction (LCR) is described. Gap filling LCR is LCR wherein at least one of the probes is recessed so that a gap is formed between the adjacent probes when they are hybridized to target. The gap is filled using polymerase and deoxyribonucleotide triphosphates before ligation of the probes together. There are single and double gap versions, depending on whether one or two probes are recessed and require filling before ligation. The improvement resides in selecting and using target sequences such that only a single type, or two types, of deoxyribonucleotide triphosphate(s) are required to fill double gaps each being 1-10 bases in length, preferably 1-3 bases. Probes having specific sequences are claimed for a number of pathogens.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TC	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark	MG	Madagascar		
ES	Spain				

-1-

5

## AMPLIFICATION OF TARGET NUCLEIC ACIDS USING GAP FILLING LIGASE CHAIN REACTION

### BACKGROUND

This invention relates to a method of performing ligase chain reaction (LCR) amplification and, particularly, to a method of ligase chain reaction amplification wherein at least one of the probes is reversibly modified at the intended point of ligation so that it is not a substrate for the ligase catalyzed reaction. Exemplary modifications include the absence of one or more nucleic acid bases to form a "recess". The modified end prevents or reduces target independent blunt-end ligation of probe duplexes and is later corrected in a target dependent manner to enable ligation.

In many cases, the feasibility of a nucleic acid based diagnostic assay is dependent on the ability to amplify the signal generated by only a few molecules of target. Although signal amplification is one potential solution, target amplification is often the preferred solution in nucleic acid based assays. Target amplification involves the repeated copying or duplication of sections of the nucleic acid designated as the target.

In the target amplification technique known as polymerase chain reaction (PCR) a pair of primers are employed in excess to hybridize at the outside ends of complementary strands of the target nucleic acid. The primers are each extended by a polymerase using the target nucleic acid as a template. The extension products become target sequences themselves, following dissociation from the original target strand. New primers are then hybridized and extended by a polymerase, and the cycle is repeated to increase geometrically the number of target sequence molecules. PCR is described further in U.S. Patents 4,683,195 and 4,683,202.

An alternate mechanism for target amplification is known as ligase chain reaction (LCR). In LCR, two primary (first and second probes) and two secondary (third and fourth) probes are employed in excess. The first probe hybridizes to a first segment of the target strand and the second probe hybridizes to a second segment of the target strand, the first and second segments being contiguous so that the primary probes abut one another in 5' phosphate-3' hydroxyl relationship and so that a ligase can covalently fuse or ligate the two probes into a fused product. In addition, a third (secondary) probe can hybridize to a portion of the first probe and a

-2-

fourth (secondary) probe can hybridize to a portion of the second probe in a similar abutting fashion. Of course, if the target is initially double stranded, the secondary probes will also hybridize to the target complement in the first instance. Once the fused strand of primary probes is separated from the target strand, it will hybridize with the third and fourth probes which can be ligated to form a complementary, secondary fused product. In order to understand LCR and the improvements described herein, it is important to realize that the fused products are functionally equivalent to either the target or its complement. By repeated cycles of hybridization and ligation, amplification of the target sequence is achieved. This technique is described more completely in EP-A-320 308 and EP-A-439 182.

A potential problem associated with ligase chain reaction is background signal caused by target independent ligation of the probes. Since the third probe hybridizes to the first probe and the fourth probe hybridizes to the second probe, the probes, which are added in excess, can easily form duplexes among themselves. These duplexes can become ligated independently of the presence of target to form a fused product which is then indistinguishable from the desired amplified target, yet which is still capable of supporting further amplification. Although target independent blunt-end ligation of these duplexes is a relatively rare event, it is sufficiently common to cause undesirably high background signals in diagnostic assays.

Some attempts to overcome this background problem have been published. For example, WO 90/01069 (Segev Diagnostics) and GB 2 225 112 A (Imperial Chemical Industries Plc.) describe versions of a ligation-based amplification scheme which includes a polymerase-mediated gap-filling step prior to ligation. In addition, EP-A-439 182 teaches variations of LCR that reduce background. One such variation involves gap filling and ligation. However, these references teach nothing about selection of the particular target sequences as taught and claimed in the present application.

It is therefore a primary object of the present invention to improve the sensitivity of nucleic acid based assays by decreasing the occurrence of target independent ligation which causes falsely positive background signal. This object is met in the present invention by modifying at least one probe end so that when hybridized with its complementary probe, the resulting duplex is not "blunt-ended" (i.e. ligatable) with respect to the partner complementary probe duplexes. After hybridizing to the target, the modified ends are "corrected" in a target dependent fashion to render the adjacent probes ligatable. Several features of the probes and the associated target sequences taught in this application make this task particularly elegant.

-3-

According to one feature, the probes have recesses relative to the point of ligation which create a gap when hybridized to the target. The gap is then filled in a target dependent manner to render the probes ligatable. Gap filling can be accomplished by extension of one or more probes, preferably two probes. Targets are selected and probes are designed so that only one, or a maximum of two, of the four deoxyribonucleotide triphosphates is needed to fill both the gaps.

## SUMMARY OF THE INVENTION

In a first aspect, the invention relates to a method for detecting the presence of a target nucleic acid sequence in a sample, said method employing the ligase chain reaction to create geometrically increasing numbers of reorganized probe molecules in the presence of said target sequence, said method comprising:

a) providing a sample suspected to contain nucleic acid, the nucleic acid having a target sequence of the formula:

15 
$$5'-(N)_h X E_p F_q Z (N)_k -3'$$

wherein E represents any base, F represents any base except E, p and q are independently integers from 1 to about 10, X represents any base except E or F, Z represents any base except F or E, N represents any base, and h and k are independently integers from 5 to 100, provided the sums (p+h) and (q+k) each are greater than 10, wherein (N)<sub>h</sub> and (N)<sub>k</sub> represent sequences that are characteristic for the target sought to be detected;

b) providing a plurality of each of four probes having the formulas:

A:  $5'-(N)_{h'} X E_p -3'$

A':  $3'-(N')_{h'} X -5'$

25 B:  $5'-Z (N)_{k'} -3'$

B':  $3'-F'_q Z' (N')_{k'} -5'$

wherein the symbols have the same meanings given above, except that h' and k' need only approximate h and k, respectively, varying by no more than 25%, and wherein at least one of the probes is labeled with a moiety capable of detection; and also providing deoxyribonucleotide triphosphates of E' and F, a polymerase reagent and a ligase reagent;

c) performing the following cycle at least once:

i) mixing said probes with said sample under hybridizing conditions to allow probes to hybridize to the target sequence and its complement if present, or to reorganized probes created therefrom;

ii) using target sequence or reorganized probes created therefrom as template, extending probe A with said polymerase reagent by adding F

-4-

deoxyribonucleotide triphosphates to its 3' end, and extending probe B' with said polymerase reagent by adding E' deoxyribonucleotide triphosphates to its 3' end

- iii) ligating extended probe A to probe B, and extended probe B' to probe A', using said ligase reagent to form reorganized probe molecules; and
- 5 iv) providing denaturing conditions to separate said reorganized probe molecules from said template;
- d) separating reorganized probe molecules from unreorganized labeled probes; and
- e) detecting the presence of said label in the reorganized or fraction as a
- 10 measure of the presence of the target sequence.

In especially preferred embodiments, F is E' so that both gaps can be filled by the deoxyribonucleotide triphosphate of only E'. Preferably p and q are small, ie., between 1 and 3, inclusive. Integers p and q may be equal or may differ by one or more. Preferably, the label comprises one or more haptens covalently bound to at

15 least one of the probes.

In another aspect, the invention relates to diagnostic kits for the detection of a target nucleic acid sequence in a sample, said nucleic acid having a target sequence comprising



- 20 wherein E and F independently represent any base, p and q are independently integers from 1 to about 10, X represents any base except E or F', Z represents any base except F or E', N represents any base, and h and k are independently integers from 5 to 100, provided the sums (p+h) and (q+k) each are greater than 10, wherein (N)<sub>h</sub> and (N)<sub>k</sub> represent sequences that are characteristic for the target sought to be
- 25 detected, said kit comprising in combination:

(a) four probes, having the formulas:

- 30 A:  $5'-(N)_h X E_p -3'$   
 A':  $3'-(N')_h X' -5'$   
 B:  $5'-Z (N)_k -3'$   
 B':  $3'-F'_q Z' (N')_k -5'$

wherein the symbols have the same meanings given above, except that h' and k' need only approximate h and k, respectively, varying by no more than 25%, and wherein at least one of the probes is labeled with a moiety capable of detection;

- (b) a polymerase reagent capable of extending probe A with said polymerase reagent in a target dependent manner by adding F deoxyribonucleotide triphosphates to its 3' end, and extending probe B' with said polymerase reagent in a target dependent manner by adding E' deoxyribonucleotide triphosphates to its 3' end, thereby to
- 35 render the primary probes ligatable to one another when hybridized to target and,

-5-

optionally, to render the secondary probes ligatable to one another when hybridized to target;

(c) deoxyribonucleotide triphosphates of E' and F; and

(d) a ligase reagent for ligating the extended probe A to probe B, and extended probe B' to probe A', thereby to form reorganized probe molecules.

Especially preferred kits, also include means for separating reorganized probe molecules from unreorganized labeled probes and/or means for detecting the detectable label.

In yet another aspect, the invention relates to compositions of matter each comprising a mixture of four probes, the four probes being selected from specific sets of sequences given in the examples. These compositions of matter have utility in detecting the presence of certain pathogens according to the methods described above. It will, of course, be realized that sequences being slightly shorter or longer than the exemplified and claimed sequences are deemed to fall within the scope of the invention, provided they are capable of hybridizing with the same locations on the targets as the claimed sequences.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a graphic representation of the process of ligase chain reaction as it is known in the prior art.

Figure 2 is a graphic representation of a generalized double gap variation of the invention.

Figure 3 is a graphic representation of a somewhat more specific double gap variation of the invention.

-6-

## DETAILED DESCRIPTION

For purposes of this invention, the target sequence is described to be single stranded. However, this should be understood to include the case where the target is actually double stranded but is simply separated from its complement prior to hybridization with the probes. In the case of double stranded target, the third and fourth (secondary) probes, A' and B', respectively, will participate in the initial step by hybridizing to the target complement. In the case of single stranded target, they will not participate in the initial hybridization step, but will participate in subsequent hybridization steps, combining with the primary fused sequence produced by ligating the first and second probes. Target sequences may comprise deoxyribonucleic acid (DNA) or ribonucleic acid (RNA).

Target sequences may be the nucleic acid of virtually any entity which contains nucleic acid. For example, but not limitation, the bacteria and viruses illustrated in the examples have nucleic acid target sequences. Target sequences may, but need not, represent pathogenic organisms or viruses. Target sequences may also represent nucleic acid which is being examined for the presence of a specific allele, or for a specific mutation of deletion, such as is often the case in genetic testing. Finally, target sequences may be simply any nucleic acid sequence of general interest, such as may be the case in forensic analysis of DNA.

Throughout this application, the "prime" (') designation is used to indicate a complementary base or sequence. A probe is "complementary" to another probe if it hybridizes to the other probe and has substantially complementary base pairs in the hybridized region. Thus, probe A can be complementary to A' even though it may have ends not coterminal with A'. The same is true of B and B'. Similarly, the short sequences  $X_n$  and  $Y_m$  have complementary sequences designated as  $X'_n$  and  $Y'_m$ , respectively. Finally, the complement of a single base, e.g. Q, is designated as Q'. As used herein with respect to sequences, "complementary" encompasses sequences that have mismatched base pairs in the hybridizable region, provided they can be made to hybridize under assay conditions.

It is also to be understood that the term "the 4 bases" shall refer to Guanine (G), Cytosine (C), Adenine (A) and Thymine (T) when the context is that of DNA; but shall refer to Guanine (G), Cytosine (C), Adenine (A) and Uracil (U) in the context of RNA. The term also includes analogs and derivatives of the bases named above. Although the degenerate base Inosine (I) may be employed with this invention, it is not preferred to use I within modified portions of the probes according to the invention.

It is an important feature of the present invention that instead of using two pairs of probes capable of forming blunt-ended duplexes, at least one probe of one of



-7-

the probe pairs initially includes a "modified" end which renders the resultant duplex "nonblunt" and/or not a suitable substrate for the ligase catalyzed fusion of the two probe duplexes. A "modified end" is defined with respect to the point of ligation rather than with respect to its complementary probe. A "modified end" has omitted  
5 bases to create a "gap" between one probe terminus and the next probe terminus (See, e.g., probes A' and B, of Figs. 2 and 3.)

By convention in this application, a modified end is referred to herein as a "recess", the recess being the gap between two primary or secondary probes after hybridizing to the target. The presence of these modified ends reduces the falsely  
10 positive signal created by blunt-end ligation of complementary probe duplexes to one another in the absence of target.

"Correction" of the modification is subsequently carried out to render the probes ligatable. As used herein "correction" refers to the process of rendering, in a target dependent manner, the two primary probes or the two secondary probes  
15 ligatable to their partners. Thus, only those probes hybridized to target, target complement or polynucleotide sequences generated therefrom are "corrected." "Correction" can be accomplished by several procedures, depending on the type of modified end used.

As used herein, "point of ligation" or "intended point of ligation" refers to a  
20 specific location between two probe partners that are to be ligated in a template-dependent manner. It is the site at which the "corrected" probe lies adjacent its partner in 3' hydroxyl-5' phosphate relationship. For each set of four LCR probes there are two "points of ligation", a point for the primary probe partners and a point for the secondary probe partners. In conventional LCR the two points of ligation are  
25 opposite one another, thus forming blunt ended duplexes when the probe pairs hybridize to one another. In the present invention, the points of ligation are not opposite one another; but are displaced from one another in the "recess" embodiments by one or more bases by virtue of the gaps. The exact point(s) of ligation varies depending on the embodiment and, thus, this term is further defined in the context of  
30 each embodiment.

Each of the probes may comprise deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). It is a routine matter to synthesize the desired probes using conventional nucleotide phosphoramidite chemistry and the instruments available from Applied Biosystems, Inc, (Foster City, CA); DuPont, (Wilmington, DE); or  
35 Milligen, (Bedford, MA). Phosphorylation of the 5' ends of the appropriate probes (eg. A' and B), while necessary for ligation by ligase, may be accomplished by a kinase or by commercial synthesis reagents, as is known in the art.

-8-

Throughout this application, the bases X, Y and Q, and their complements are described as being selected from certain subsets (N or M) of the 4 bases. In reality, the sequences are not "selected" at all, but are dictated by the sequence of the target strand. The term "selected" in this context is taken to mean that a target sequence  
5 having the desired characteristics is located and probes are constructed around an appropriate segment(s) of the target sequence.

In general, the methods of the invention comprise repeated steps of (a) hybridizing the modified probes to the target (and, if double stranded so that target complement is present, to the target complement); (b) correcting the modification in  
10 a target dependent manner to render the probes ligatable; (c) ligating the corrected probe to its partner to form a fused or ligated product; and (d) dissociating the fused product from the target and repeating the hybridization, correction and ligation steps to amplify the desired target sequence. Steps (a), (c) and (d) are essentially the same for all of the embodiments and can be discussed together. They are generally the  
15 same steps that one would employ in conventional LCR. Step (b) varies depending on the type of modification employed and each different type is discussed separately.

Hybridization of probes to target (and optionally to target complement) is adequately explained in the prior art; e.g EP-320 308. Probe length, probe concentration and stringency of conditions all affect the degree and rate at which  
20 hybridization will occur. Preferably, the probes are sufficiently long to provide the desired specificity; i.e. to avoid being hybridizable to nontarget sequences in the sample. Typically, probes on the order of 15 to 100 bases serve this purpose. Presently preferred are probes having a length of from about 15 to about 40 bases.

The probes are added in approximately equimolar concentration since they are  
25 expected to react stoichiometrically. Each probe is present in a concentration ranging from about 5 nanomolar (nM) to about 90 nM; preferably from about 10 nM to about 35 nM. For a standard reaction volume of 50  $\mu$ L, this is equivalent to adding from about  $3 \times 10^{11}$  to about  $1 \times 10^{12}$  molecules of each probe; and around  $5 \times 10^{11}$  molecules per 50  $\mu$ L has been a good starting point. The optimum quantity of probe  
30 used for each reaction also varies depending on the number of cycles which must be performed and, of course, the reaction volume. Probe concentrations can readily be determined by one of ordinary skill in this art to provide optimum signal for a given number of cycles.

The stringency of conditions is generally known to those in the art to be  
35 dependent on temperature, solvent and other parameters. Perhaps the most easily controlled of these parameters is temperature and thus it is generally the stringency parameter varied in the performance of LCR. Since the stringency conditions required for practicing this invention are not unlike those of ordinary LCR, further

detail is deemed unnecessary, the routine practitioner being guided by the examples which follow.

The next step in the general method follows the specific correction step and comprises the ligation of one probe to its adjacent partner. Thus, each corrected  
5 primary probe is ligated to its associated primary probe and each corrected secondary probe is ligated to its associated secondary probe. An "adjacent" probe is either one of two probes hybridizable with the target in a contiguous orientation, one of which lies with its phosphorylated 5' end in abutment with the 3' hydroxyl end of the partner probe. "Adjacent" probes are created upon correction of the modified end(s) in a  
10 target dependent manner. Since enzymatic ligation is the preferred method of covalently attaching two adjacent probes, the term "ligation" will be used throughout the application. However, "ligation" is a general term and is to be understood to include any method of covalently attaching two probes. One alternative to enzymatic ligation is photo-ligation as described in EP-A-324 616.

15 The conditions and reagents which make possible the preferred enzymatic ligation step are generally known to those of ordinary skill in the art and are disclosed in the references mentioned in background. Ligating reagents useful in the present invention include T4 ligase, and prokaryotic ligases such as *E coli* ligase, and *Thermus thermophilus* ligase (e.g., ATCC 27634) as taught in EP-320 308. This  
20 latter ligase is presently preferred for its ability to maintain activity during the thermal cycling of LCR. Absent a thermally stable ligase, the ligase must be added again each time the cycle is repeated. Also useful are eukaryotic ligases, including DNA ligase of *Drosophila*, reported by Rabin, et al., *J. Biol. Chem.* 261:10637-10647 (1986).

25 Once ligated, the fused, reorganized probe is dissociated (e.g. melted) from the target and, as with conventional LCR, the process is repeated for several cycles. The number of repeat cycles may vary from 1 to about 100, although from about 15 to about 70 are preferred presently.

It is desirable to design probes so that when hybridized to their  
30 complementary (secondary) probes, the ends away from the point of intended ligation are not able themselves to participate in other unwanted ligation reactions. Thus, ligatable sticky or blunt ends should be avoided. If such ends must be used, then 5' terminal phosphates should be avoided, eliminated or blocked. This can be accomplished either through synthesizing oligonucleotide probes (which normally  
35 carry no 5' terminal phosphate groups), or through the use of phosphatase enzymes to remove terminal phosphates (e.g. from oligonucleotides generated through restriction digests of DNA). Alternatively, ligation of the "wrong" outside ends of the probes can be prevented by blocking the end of at least one of the probes with a "hook"

-10-

or marker moiety as will be described in detail below. In the absence of one of the above techniques, the outside ends of the probes can be staggered so that if they are joined, they will not serve as template for exponential amplification.

Following amplification, the amplified sequences can be detected by a number of conventional ways known in the art. Typically, detection is performed after separation, by determining the amount of label in the separated fraction. Of course, label in the separated fraction can also be determined subtractively by knowing the total amount of label added to the system and measuring the amount present in the unseparated fraction. Separation may be accomplished by electrophoresis, by chromatography or by the preferred method described below.

In a particularly preferred configuration, haptens, or "hooks", are attached at the available outside ends of at least two probes (opposite ends of fused product), and preferably to the outside ends of all four probes. A "hook" is any moiety having a specific ligand-receptor affinity. Typically, the hook(s) at one end of the fused product (e.g. the 5' end of A and the 3' end of A') comprises an antigen or hapten capable of being immobilized by a specific binding reagent (such as antibody or avidin) coated onto a solid phase. The hook(s) at the other end (e.g. the 3' end of B and the 5' end of B') contains a different antigen or hapten capable of being recognized by a label or a label system such as an antibody-enzyme conjugate. Exemplary hooks include biotin, fluorescein and digoxin among many others known in the art. A substrate is then added which is converted by the enzyme to a detectable product. EP-A-330 221 to Enzo describes oligonucleotides having a biotin molecule attached at one end.

#### 25 ENDS MODIFIED BY RECESSES

In this embodiment, modified ends are created by eliminating from one or more of the probes a short sequence of bases, thereby leaving a recess or gap between the 5' end of one probe and the 3' end of the other probe when they are both hybridized to the target (or target complement, or polynucleotide generated therefrom). In order for LCR to amplify the target, the gaps between the probes must be filled in (i.e., the modification must be "corrected"). In a first version, this can be done using a polymerase or a reverse transcriptase and an excess of deoxyribonucleotide triphosphates which are complementary to the target strand opposite the gap.

However, prior to discussing this embodiment in detail, a brief digression on set terminology may be helpful. A set (e.g., S) consists of all the elements contained within the set S. The set "not S" then consists of all the remaining elements of the "Universe" which are not found in S. The "Universe" for purposes of this application consists of the four bases G, C, A and T, or G, C, A and U as described above. The

- 11 -

Intersection of set S and another set (e.g., R) consists only of those elements which are found in both S and R. Thus, as used in this application, the set "not N and not M" consists of those bases which are present in *neither* the gap  $X_n$  nor the gap  $Y_m$ .

- According to this invention, the set "not N and not M" must not be an empty set; i.e. at least one base must remain in this set to code for the "stopbase".

#### Gap Filling by Extension:

- In accordance with this first version, the invention involves repeated steps of (a) hybridizing the probes to the target (and, if double stranded so that target complement is present, to the target complement); (b) extending at least one probe to fill in at least one gap, designated  $X_n$ ; (c) ligating the extended probe to the adjacent probe to form a fused or ligated product; and (d) dissociating the fused product from the target and repeating the hybridization, extension and ligation steps to amplify the desired target sequence.

- In this version, which includes both single gap ("SG") and double gap ("DG") configurations, the "gaps"  $X_n$  and  $Y_m$  which impart the "modified ends" are "corrected" by extending one or both of the modified probes using a polymerase or a reverse transcriptase. Generally, extension of a probe hybridized to a DNA target is accomplished by a DNA polymerase or a Klenow fragment as is known in the art. In the case of an RNA target, extension is accomplished by a reverse transcriptase.
- Exemplary reverse transcriptases include those from avian myeloblastosis virus (AMV) and Moloney murine leukemia virus (M-MuLV) generally available to those skilled in the art. Certain DNA polymerases will also recognize RNA as template under certain conditions. It is, of course, preferable to utilize extension reagents which are thermally stable and can withstand the cycling of high temperatures required for LCR. If the extension reagent is not thermally stable, it typically must be re-added at each cycle of LCR. Such thermostable polymerases presently include AmpliTaq<sup>TM</sup>, (available from Cetus-Perkin Elmer), *Thermus* polymerase (available from Molecular Biology Resources, Inc. Milwaukee, WI, "MBR") and recombinant or purified native polymerases from *Thermus aquaticus*, *Thermus thermophilus* or other species known to be thermostable.

- Correction by extension in this manner requires the presence in the reaction mixture of deoxyribonucleotide triphosphates (dNTP's) complementary to the bases of the target in the gap region(s). More specifically, for a gap having the sequence  $X_n$ , the dNTP's that must be supplied are designated dXTP wherein X' stands for the complements of each base in the gap  $X_n$ . The dNTP's are commercially available from a number of sources, including Pharmacia (Piscataway, NJ) and Bethesda Research Laboratories (Gaithersburg, MD).

-12-

Extension must be terminated precisely at the point of ligation so that the extended probe abuts the adjacent probe and can be ligated to it. "Stopbases" are employed for this purpose, (See figure 2). A "stopbase", designated Q', is defined in terms of its complement, Q and is accomplished by omitting from the reaction mixture, dNTP's that are complementary to Q; i.e. by omitting dQ'TP from the reaction mixture. Thus it is seen how the bases for the gap sequence(s) must be selected from a set, N, consisting of only three of the four bases, so that the complementary three of the four dNTP's are added to the reaction mixture. When the fourth dNTP, dQ'TP, is absent from the reaction mixture extension will terminate at the desired point of ligation. It follows that Q' is the first base in the adjacent probe, and the base on the target which codes for the stopbase is the first base adjacent the gap (on the 5' end of the  $X_n$  gap in Figure 2).

While the concept is easiest to grasp in the SG configuration, it should be understood that the SG variation is merely a special case of the double gap (DG) variation discussed below (in SG,  $m=0$ ). However, only the DG configuration is relevant to the presently claimed invention. As shown in Fig 2, a first probe, A, hybridizes to a first segment of the target strand, T. A second probe, B, hybridizes to a second segment of the target strand, leaving a gap of one or more bases between the two probes. This gap is designated  $X_n$  on the target. Following extension and ligation, a third probe, A', is hybridizable to a first portion of reorganized probe, A:B; and a fourth probe, B', is hybridizable to a second portion of reorganized probe, A:B. In the DG configuration, the third probe, A', and the fourth probe, B', hybridize such that a gap of one or more bases lies between the probes. This gap is designated  $Y_m$  on the reorganized probe (and on target complement). As is shown in Fig. 2, the target strand, T, may be double stranded, having a complement, T'. In this case, the third and fourth probes may participate in the initial hybridization by hybridizing to first and second segments of the target complement.

Extension by polymerase or transcriptase proceeds in a 5' to 3' direction. Consequently, the 3' ends of both A and B' will be extendable by polymerase in the absence of anything to prevent extension. Extension is terminated when the next base called for by the template is absent from the reaction mixture. Thus, probe A is extended through gap  $X_n$  until stopbase complement (Q) is encountered along the target strand. Similarly, probe B' is extended through gap  $Y_m$  until stopbase complement (Q) is encountered (either on the target complement or on the A half of reorganized A:B). Neither probe A' nor B will serve as a template for extension of A or B', so probes A and B' are extended only if hybridized to the target (or to reorganized polynucleotide products from previous cycles).

- 13 -

As alluded to above, it is important to terminate the extension of A and B' at the end of the respective gaps (i.e., at the point of ligation) so that the extended probe can be ligated to the 5' end of the adjacent probes, B and A'. Therefore, the reaction mixture omits the deoxyribonucleotide triphosphate complementary to the base (Q) immediately adjacent the 5' end of gaps  $X_n$  and  $Y_m$ . Of course, it will be understood that it is not required that the same base stop extension in both directions. A different base can be used provided it is not needed to fill either of the gaps. It should now be apparent that the actual points of ligation in this embodiment are always at the 5' ends of probes A' and B. It is not by mere coincidence that these are also the locations of the stopbases Q'.

Accordingly, the gaps  $X_n$  and  $Y_m$  can be any number of bases long, i.e., n can be any integer greater than or equal to 1, and m is any integer greater than 0. It is to be realized, however, that the choice of which gap is  $X_n$  and which is  $Y_m$  is arbitrary in the first place; but n and m cannot both be zero. The gaps need not be the same length, i.e., m need not equal n. When, m equals zero, the double gap variation degenerates into the specialized case of the single gap, which is not used in the embodiment being claimed herein. The only restriction on the bases X is that they be selected from a set N which consists of from 1 to any 3 of the four bases. Similarly, the bases Y are drawn from set M. Since at least one stopbase Q' must be maintained, the combined sets N and M which represent the possible bases for X and Y, respectively, must include no more than three of the four bases. Accordingly, Y can be from zero to any three of the four bases provided that at least one base remains in the set "not N and not M". If set N constitutes less than three of the four bases, then Y can be a base that is not within N so long as there is at least one base remaining, the complement of which can serve as the stopbase Q' for termination of probe extension. A single stopbase can serve to terminate extension in both the  $X_n$  and  $Y_m$  gaps.

A second limitation on sequence  $Y_m$  occurs if m equals n. If the gaps are the same length, the sequence  $Y_m$  should not be complementary to the sequence  $X_n$  or the 3' ends of probes A and B' would constitute "sticky ends". "Sticky ends" would permit a target independent double stranded complex to form wherein probe A hybridizes to probe B' such that ligations and amplification would proceed. Rather, when m equals n it is preferred that  $Y_m$  not be complementary to  $X_n$ . In other words, the ends of probes A and B' should at least be "slippery ends" which may be the same length, but are not complementary.

In a preferred aspect of the invention, the fourth probe B' includes a 3' terminal sequence of  $X_n$ , identical in length to the  $X_n$  sequence gap in the target. This arrangement is not essential to the invention, however, as the gap need only be formed between the probes. Thus, the 3' terminus of the fourth probe B' may stop short of

- 14 -

the 3' end of sequence  $X_n$ , provided there is no 3' recessed end with respect to the second probe B. Since extension occurs in a 5' to 3' direction and dX'TPs must be present anyway (to extend through  $X_n$ ), probe B' would be extended through the gap, (both  $Y_m$  and any remainder of  $X_n$ ) just as the first probe A is extended through the

5  $X_n$  gap.

The general method of the invention employing the double gap embodiment is now described briefly. First, probes A, and B are allowed to hybridize to target, if present. If the target is double stranded, it is first denatured and probes A' and B' can

10 also hybridize to the target complement. In the presence of a suitable polymerase, probes A and B' are extended by the addition of dX'TPs and dY'TPs, respectively, to their 3' ends using the target strand as a template. However, when the polymerase encounters the base Q on the template, it is unable further to extend probe A since dQ'TP is not provided in the reaction mixture as dNTP's. Extension terminates

15 precisely at the point of ligation with the extended 3' end of probe A abutting the 5' end of probe B. It is not known if the presence of probes A' and B on the template will terminate extension. But even in their absence, extension is terminated if attention is paid to proper stopbase selection.

Next, a ligase is employed to join the 3' hydroxyl end of extended probe A to

20 the 5' phosphate end of probe B to form a double stranded complex of fused or ligated primary probe and target strand. If the target is double stranded and has a complement T', the ligase will also join probes A' and B' in the initial cycle if they are hybridized to the target complement. If they are hybridized to excess probes A and B rather than target complement, ligation is inhibited since the ends are neither blunt

25 nor sticky and there is no substrate for ligation.

Subsequently, the double stranded complexes are dissociated and new probes A, A', B and B' are permitted to hybridize to the target, the target complement, and both of the fused polynucleotides from the first cycle. Extension and ligation occur as before and the process can be repeated.

30 Some exemplary combinations of X's and Y's, their dNTP counterparts and the resultant possibilities for Q and Q' are given in Table I.



-15-

TABLE I

ILLUSTRATIVE GAP SEQUENCES, REQUIRED dNTPs, and POSSIBLE  
COMBINATIONS FOR Q and Q' IN DOUBLE GAP VARIATION

<u>X<sub>n</sub>/N</u>	<u>Y<sub>m</sub>/M</u>	<u>X'TPs</u>	<u>Y'TPs</u>	<u>not N and not M*</u>	<u>STOPBASE Q'</u>
A	A	T	T	T, C, G	A, G, C
G	T	C	A	C, A	G, T
AT	AT	T, A	T, A	C, G	C, G
AC	GA	T, G	C, T	T	A
ATG	AAA	T, A, C	T	C	G
G G C C	AAACG	C, G	T, G, C	T	A
ATTGA	AGGT	T, A, C	T, C, A	C	G
C G C	G C G	complement. not permitted			

5

\*The set not N and not M provides the possible complements (Q) for the stopbase Q'.  
The actual stopbase (Q') possibilities are given in the next column.

In the case of double gaps, the length of gaps X<sub>n</sub> and Y<sub>m</sub> may be one or any  
10 integer greater than one. For example, gaps of from 1 to 20 bases would be possible.  
Practically, however, gaps of much shorter length are preferred; for example from 1  
to 3 or 5 bases. It has been found that gaps of just one or two bases greatly increase  
the ratio of true signal to background and leave the largest number of options for  
stopbases and dXTP's. Since probes are actually designed around existing targets,  
15 rather than "selecting" stopbases, single or two base gaps are desirable in most cases.  
Furthermore, it has been discovered that the double gap embodiment is preferred to  
the single gap version.

#### TARGET SEQUENCE CHARACTERISTICS

20 For clarity, double gaps are represented herein as "DG p,q" where "p" is the  
number of bases in the gap in one strand, and "q" is the number of bases in the gap of  
the other strand. Thus, a preferred double gap embodiment has two bases missing  
from each of the two probes whose 5' end participates in the ligation, and is  
designated DG 2,2. In this preferred embodiment, the 3' ends of the other two probes  
25 do not overlap; rather they terminate at the same point on the target strand (and its  
complement). Also for clarity in this application the use of a dot or period "."

- 16 -

between bases in a target sequence represents the point at which the probes, when hybridized to their complementary probes, would be blunt-ended but for the gaps in two of the probes. In other words, probes A and B' (see Fig. 3) have 3' termini which, absent the gaps in probes A' and B, would be blunt-ended at the point of ligation. It is this point at which ligation would have occurred that is designated by a dot. However, with gaps present according to the invention, probes A and B' are extended, thus creating two, offset points of actual ligation.

In a particularly preferred DG embodiment both the gaps are fillable with just one type of deoxyribonucleotide triphosphate. This requires selecting a target sequence having the sequence  $-E_p.E'_q-$ , where E represents any base, and p and q are integers between 1 and about 10. The probe sets must be designed so that the two probes whose 3' ends participate in the extension/ligation (eg. probes A and B' in Fig. 3) both terminate at the dot "." between E and E'. In order to provide a "stopbase" a further requirement is added. The target sequence must have a base other than E in the next position outward from the E string; and, similarly, must have a base other than E' in the next position outward from the E' string. This can be represented as  $-LE_p.E'_qJ-$ , where L is any base other than E and J is any base other than E'. Finally, to provide sufficient probe length for easy hybridization and, optionally, to provide specificity, the regions  $(N)_h$  and  $(N)_k$  are employed, to give:

20  $5'-(N)_hLE_p.E'_qJ(N)_k-3'$

where N represents any base independently of any other base, N; and h and k are each integers greater than 5. Generally, to provide the desired specificity, the individual probes must be at least 10 or 11 nucleotides long, i.e., the sums  $(p+h)$  and  $(q+k)$  must be at least 10; more usually will be about 20 to 40, but can be up to 100 or more. As mentioned, p and q may be the same or different, and generally will be small, eg. 1-3.

In this preferred DG 2,2 system, probes would be designed and synthesized as follows, using the designations given in Fig. 3:

30 

A'	Hapten <sub>1</sub> -3'-(N') <sub>h</sub> L'-5'	3'-E <sub>q</sub> J'(N') <sub>k</sub> -5'-Hapten <sub>2</sub>	B'
A	Hapten <sub>1</sub> -5'-(N) <sub>h</sub> LE <sub>p</sub> -3'	5'-J(N) <sub>k</sub> -5'-Hapten <sub>2</sub>	B

wherein the terms are defined as above except that h' and k' need not equal h and k exactly. It is entirely permissible for any of the probes to be slightly shorter or longer than either the target sequence or the complementary probe. This variation in length is indicated by the "'" symbol. The variation is usually no more than 20-25% of the probe length, preferably 10% or less. Of course, no variation in length is also possible. Moreover, h' and k' may represent a different number at each

- 17 -

occurrence, provided they are within the 20- 25% variation limit; thus, probe A need not be exactly p bases longer than probe A' (their (N)<sub>h</sub> portions may vary in length too) and the same is true for k' and probes B and B'.

The single deoxyribonucleotide triphosphate of E' only is needed to fill all the gaps. The bases L and J perform the same function as Q, i.e. to code for the stopbase Q'. However, designations L and J are used here because of the additional restrictions placed on them in this embodiment. As described elsewhere in this specification, haptens 1 and 2 serve to enable separation and detection of ligated products. They are particularly useful for automated detection in the IMx® instrument (Abbott Laboratories). Hapten 1 should be different than hapten 2. Virtually any low molecular weight compound to which antibodies can be raised may serve as hapten.

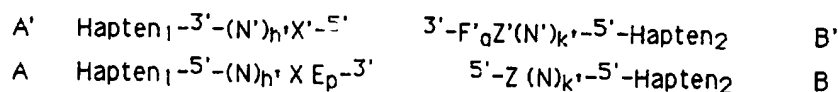
Several examples of this preferred embodiment are given in the first column of Table II below and in the Examples which follow.

Another more generalized embodiment of the invention utilizes target sequences and probe sets wherein just two types of deoxyribonucleotide triphosphates are needed to fill all the gaps. Examples of this embodiment are given in the second column of Table II and its target sequence is represented by the following formula:



wherein N, p, q, h and k are all defined as before. E may be any base as before. F may be any base except E (if it were E the ends would be "sticky" when p=q), but when it is E' this case degenerates to the simpler embodiment described above. Because of the additional flexibility in the choice of F, additional restrictions are placed on the stop bases X and Z. X may be any base but E or F', while Z may be any base but F or E'. The gaps are filled by the deoxyribonucleotide triphosphates of E' and F.

In this DG 2,2 system, probes would be designed and synthesized as follows, using the designations given in Fig. 3:



30

The two deoxyribonucleotide triphosphates needed to fill all the gaps are F and E'. The bases X and Z perform the same function as Q to code for the stop base Q'. However, designations X and Z are used here because of the additional restrictions placed on them in this embodiment. The restrictions are also different than the restrictions on L and J. The haptens are as described above.

35

It is important the "X" which codes for a stopbase here (or X or X' used in a probe) should not be confused with the X<sub>n</sub> gap or the dX'TPs needed to fill the X<sub>n</sub> gap. Regrettably, there are only twenty-six letters in the alphabet and most have been

-18-

used in this application. It is believed that confusion between the two uses of X and X' is avoided by the context.

Other variations of this invention include DG configurations other than 2,2. As shown in Table II and in the examples, DG 1,1; DG 1,2 (as well as 2,1); DG 2,3 (as well as 3,2) and DG 3,3 are also possible. In addition, gaps of up to 10 bases are included in the generic p and q case, although Table III below shows the low probability of gap junctions longer than about 3. There may be some advantage to having p be approximately equal to q (i.e. varying by only 1 or 2), but this is not a strict requirement. Though not tested, there is no reason to believe a double gap of 1,3 or 2,4 will not also work in the invention.

The above discussion describes in detail the "Special" symbols used in Table II. The "Conventional" symbols will not need further explanation to those of ordinary skill in the art.

TABLE II  
EXEMPLARY SINGLE and TWO dNTP FILL  
DOUBLE GAP ("DG") JUNCTIONS

(The dot "." serves only to align sequences in the Table and to divide between right and left probe sets: All targets are written with their 5' end to the left.)

	Gaps fillable with just one deoxyribonucleotide triphosphate type		Gaps fillable with just two deoxyribonucleotide triphosphates types	
DG 1,1 Generic:	(N) <sub>h</sub> LE.E'J(N) <sub>k</sub> E' fills (in all generic cases)		(N) <sub>h</sub> XE.FZ(N) <sub>k</sub> E' & F fill (in all generic cases)	
Subgeneric:	(N) <sub>h</sub> DC.GH(N) <sub>k</sub>	G fills	(N) <sub>h</sub> KC.TM(N) <sub>k</sub>	G&T fill
	(N) <sub>h</sub> HG.CD(N) <sub>k</sub>	C fills	(N) <sub>h</sub> YG.TR(N) <sub>k</sub>	C&T fill
	(N) <sub>h</sub> BA.TV(N) <sub>k</sub>	T fills	(N) <sub>h</sub> RC.AY(N) <sub>k</sub>	G&A fill
	(N) <sub>h</sub> VT.AB(N) <sub>k</sub>	A fills	(N) <sub>h</sub> MG.AK(N) <sub>k</sub>	C&A fill

- 19 -

DG 1,2 and 2,1	(N) <sub>h</sub> LE.E'E'J(N) <sub>k</sub>	(N) <sub>h</sub> XE.FFZ(N) <sub>k</sub>
Subgeneric:	(N) <sub>h</sub> DC.GGH(N) <sub>k</sub> G f111s (N) <sub>h</sub> HG.CCD(N) <sub>k</sub> C f111s (N) <sub>h</sub> BA.TTV(N) <sub>k</sub> T f111s (N) <sub>h</sub> VT.AAB(N) <sub>k</sub> A f111s	(N) <sub>h</sub> KC.TTM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> KCC.TM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> YG.TTR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> YGG.TR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> RC.AAY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> RCC.AY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> MG.AAK(N) <sub>k</sub> C&A f111 (N) <sub>h</sub> MGG.AK(N) <sub>k</sub> C&A f111
DG 2,2 Generic:	(N) <sub>h</sub> LEE.E'E'J(N) <sub>k</sub>	(N) <sub>h</sub> XEE.FFZ(N) <sub>k</sub>
Subgeneric:	(N) <sub>h</sub> DCC.GGH(N) <sub>k</sub> G f111s (N) <sub>h</sub> HGG.CCD(N) <sub>k</sub> C f111s (N) <sub>h</sub> BAA.TTV(N) <sub>k</sub> T f111s (N) <sub>h</sub> VTT.AAB(N) <sub>k</sub> A f111s	(N) <sub>h</sub> KCC.TTM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> YGG.TTR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> RCC.AAY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> MGG.AAK(N) <sub>k</sub> C&A f111
DG 2,3 and 3,2	(N) <sub>h</sub> LEE.E'E'E'J(N) <sub>k</sub>	(N) <sub>h</sub> XEE.FFFZ(N) <sub>k</sub>
Subgeneric:	(N) <sub>h</sub> DCCC.GGH(N) <sub>k</sub> G f111s (N) <sub>h</sub> HGGG.CCD(N) <sub>k</sub> C f111s (N) <sub>h</sub> BAAA.TTV(N) <sub>k</sub> T f111s (N) <sub>h</sub> VTTT.AAB(N) <sub>k</sub> A f111s	(N) <sub>h</sub> KCCC.TTM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> KCC.TTTM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> YGGG.TTR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> YGG.TTTR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> RCCC.AAY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> RCC.AAAY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> MGGG.AAK(N) <sub>k</sub> C&A f111 (N) <sub>h</sub> MGG.AAAK(N) <sub>k</sub> C&A f111
DG 3,3 Generic:	(N) <sub>h</sub> LEEE.E'E'E'J(N) <sub>k</sub>	(N) <sub>h</sub> XEEE.FFFZ(N) <sub>k</sub>
Subgeneric:	(N) <sub>h</sub> DCCC.GGGH(N) <sub>k</sub> G f111s (N) <sub>h</sub> HGGG.CCCD(N) <sub>k</sub> C f111s (N) <sub>h</sub> BAAA.TTTV(N) <sub>k</sub> T f111s (N) <sub>h</sub> VTTT.AAAB(N) <sub>k</sub> A f111s	(N) <sub>h</sub> KCCC.TTTM(N) <sub>k</sub> G&T f111 (N) <sub>h</sub> YGGG.TTTR(N) <sub>k</sub> C&T f111 (N) <sub>h</sub> RCCC.AAAY(N) <sub>k</sub> G&A f111 (N) <sub>h</sub> MGGG.AAAK(N) <sub>k</sub> C&A f111
DG p,q	(N) <sub>h</sub> LE <sub>p</sub> .E'qJ(N) <sub>k</sub>	(N) <sub>h</sub> XE <sub>p</sub> .FqZ(N) <sub>k</sub>
GENERAL CASE	E' alone f111s	E' and F alone f111

-20-

Where the symbols have the following meanings:

CONVENTIONAL		SPECIAL	
A	Adenine	E	any base
B	any base but adenine (not A)	F	any base
C	Cytosine	L	any base except E
D	any base but cytosine (not C)	J	any base except E'
G	Guanine	X	any base except E or F'
H	any base but guanine (not G)	Z	any base except F or E'
K	G or T/U only	h, k	any integer greater than 1
M	A or C only	p, q	any integer between 1 and about 10
N	any base		
R	A or G only		
S	C or G only		
T	Thymine		
U	Uracil		
V	any base but thymine/uracil		
W	A or T/U only		
Y	C or T/U only		

It will be understood that those of ordinary skill in the art will know how to search for and identify specific target sequences meeting the requirements outlined above. For example many databases contain sequence information (eg. GENBANK, NBRF, EMBL or Swiss-Prot) which can be searched by readily available computer software (eg. MacVector, MacMolly or Geneworks). It will also be realized that in any organism's known genome, multiple locations meeting the requirements will generally be found. For example, searching just the SF2CG isolate of HIV 1 (see Examples 1-5), which contains approximately 9.7 kilobases, reveals over 2000 possible DG locations. The breakdown for each type of gap and fill situation is given in Table III. It should be realized that when  $p=q$ , a DG  $p,q$  on one strand at one location also comprises a DG  $p,q$  site on the opposite strand. In the situation where  $F=E'$  (ie. a single dNTP fill) the two target sites are really just one location. Thus, for all symmetric double gaps (where  $p=q$ ) the Table reports half the number actually found by searching. In contrast, an asymmetric DG (eg, a DG 1,2 or DG 2,3) on one strand represents a second, different site (but DG 2,1 or DG 3,2, this time) on the opposite strand. Thus, asymmetric sites are reported in the Table without halving the number.

-21-

TABLE III  
Representative Numbers of Gaps by type  
in an Organism's Genomic DNA

Type of GAP	Gaps fillable with just one deoxyribo-nucleotide triphosphate type	Gaps fillable with just two deoxyribo-nucleotide triphosphates types
HIV 1 (SF 2CG Isolate) <sup>1</sup> (9.7 kb)		
DG 1,1	1053	1099
DG 1,2	538	636
DG 2,2	75	71
DG 2,3	42	43
DG 3,3	8	5
<i>C. trachomatis</i> MOMP <sup>2</sup> (2.2 kb)		
DG 1,2	120	136
DG 2,2	24	18
DG 2,3	8	10
DG 3,3	0	0
<i>C. trachomatis</i> cryptic plasmid <sup>3</sup> (7.5 kb)		
DG 1,2	415	488
DG 2,2	61	55
DG 2,3	39	24
DG 3,3	6	6

1. Sanchez-Pescado, R., et al. *Science* 227:484-492 (1985), see examples 1-5.
2. Baehr, W. et al. *Proc. Natl. Acad. Sci. USA* 85:4000-4004 (1988) [EMBL Genbank Data Base: IntelliGenetics, Accession #J03813], see examples 11-12.
3. Hatt, C. et al. *Nuc. Acids Res.* 16:4053-4067 (1988), see example 13.

For the HIV 1 isolate, thirteen of the 75 DG 2,2 single dNTP fill sites are located within the *gag* region. These 13 locations are listed in Appendix A.

In cases where genomic sequences are not known, at least certain portions must be established prior to using any version of LCR. Typically, one of ordinary skill can determine a sequence through routine methods, such as cloning and sequencing as taught in conventional textbooks on the subject, such as Maniatis, T., et al. *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982 and 1989).

Whether one is seeking a universal probe set for several types, strains, or serovars, etc. (as might be the case for HIV or *Chlamydia*) or one is seeking a

-22-

specific probe set to identify just one type (as might be the case for HPV), one of ordinary skill will know how to compare sequences obtained from such a search with the genomes of other organisms to identify conserved or unique sequences, respectively. In addition, routine laboratory investigations may be needed to confirm the universality or specificity of a particular probe set. The length and nature of (N)<sub>h</sub> and (N)<sub>k</sub> may greatly affect the universality or specificity of the respective probes, although it is also possible to design probes such that specificity is gained by the gap regions.

#### 10        Further Features

In a variation of either "recessed" embodiment, the deoxyribonucleotide triphosphates used in filling the gaps may be modified to contain a marker moiety. Exemplary markers include direct labels such as radioisotopes, or hooks such as biotin, digoxin or other hapten capable of being recognized by either a solid phase or a label producing system. Isotopic labels include <sup>32</sup>P, and <sup>3</sup>H among others.

15        Incorporation of the marker into the dNTP's is generally a matter of conventional organic chemistry. Linkers or spacers may be used but are not essential. It is only important that the modified dNTP be able to be incorporated into the gap opposite its complement on the target strand and be covalently bonded to the adjacent base.

#### Examples

The invention will now be described further by way of examples which are illustrative of the invention and are not intended to limit it in any way. For example, sequences of specific length are listed. It should be understood that sequences covering the same map positions but having slightly fewer or greater numbers of bases are deemed to be equivalents of these sequences and fall within the scope of the invention, provided they will hybridize to the same positions on the target as the listed sequences.

30        Unless otherwise indicated, sequences in the examples are written with the 5' end toward the left. In the examples, some probes are labeled with haptens: "FL" represents a fluorescein hapten and "BIO" represents a biotin hapten, covalently attached to the probe. The haptens are attached at either the 3' or 5' terminus as shown. Throughout the examples, the numbering system for probes corresponds to a map position followed by a decimal .1, .2, .3, or .4, which correspond to probes A, A', B, B', respectively, of Fig. 3. Thus, arranging the probes as:

.1 : .3  
.2 : .4



-23-

yields an arrangement which is upside-down when compared to Fig. 2 or 3. The map position given is for the left-most base when the probes are aligned as if hybridized. The remaining map positions can be calculated from this information, knowing that the left-most position bears the lowest number.

5 In most of the examples, results were read in an IMx® instrument. This is commercially available from Abbott Laboratories and is described in EP-A-288 793 and in Fiore, M. et al *Clin. Chem.*, 34/9:1726-1732 (1988). It should be noted that the IMx instrument typically generates "machine" noise or background in the range of 2-8 counts/sec/sec.

10 Quantities of polymerase are expressed in units, defined (eg., by MBR) as follows: 1 unit of enzyme equals the amount of enzyme required to incorporate 10 nanomoles total nucleotides into acid-insoluble material in 30 min at 70°C. Units of ligase enzyme are defined (internally by Abbott Laboratories) as: 1 mg of 95% purified *Thermus thermophilus* DNA ligase has a specific activity of about  $1 \times 10^8$   
15 units. While this is not precisely standardized and may vary by as much as 20%, optimization is within the skill of the routine practitioner.

Examples 1-12 relate to detection of Human Immunodeficiency Virus (HIV1) using several different sequences. The sequences, their map position (SF2CG isolate  
20 according to Sanchez-Pescado, R., et al. *Science* 227:484-492 (1985)), and their Sequence ID Nos. are given in Table IV, below. Appendix A gives the number and map positions of DG2,2 targets in the *gag* region of HIV1.

Table IV.

25	MAP POS	SEQUENCE	SEQ ID No.
	1667.1:	FL-AACCCCTTTAG AGACTATGTA GACC	1
	1667.2:	TCTACA AGT CTCTAAAGGG TTC-FL	2
30	1667.3:	TTCTATAAAA CTCTAAGAGC CGA-BIO	3
	1667.4:	BIO-CGGCTCTTAG AGTTTTATAG AACC	4
	912.1:	FL-AGAACGATTC GCAGTCAATC CTGG	5
	912.2:	AGGATTGACT GCGAATCGTT CTA-FL	6
35	912.3:	TGTTAGAAAC ATCAGAAGGC TGC-BIO	7
	912.4:	BIO-CAGCCTTCTG ATGTTTCTAA CAGG	8
	2086.1:	FL-GCTAATTTTT TAGGGAAGAT CTGG	9
	2086.2:	AGATCTTCCC TAAAAAATTA GCC-FL	10
40	2086.3:	TTCTACAAG GGAACGCCAG GGA-BIO	11
	2086.4:	BIO-CCCTGGCGTT CCCTTG TAGG AAGG	12

-24-

	789.1:	FL-GATGGGTGCG AGAGCGTCGG TATT	13
	789.2:	TACCGACGCT CTCGCACCCA TCT-FL	14
	789.3:	GCGGGGAGA ATTAGATAAA TGG-BIO	15
	789.4:	BIO-CATTTATCTA ATTCTCCCC GCTT	16
5	508.1:	FL-ACCCACTGCT TAAGCCTCAA TAAAG	17
	508.2:	TTTATTGAGG CTTAAGCAGT GGGTT-FL	18
	508.3:	TTGCCTTGAG TGCTTCAAGT AGTGT-BIO	19
	508.4:	BIO-CACTACTTGA AGCACTCAAG GCAAG	20
10	5569.1:	FL-GAACAAGCCC CAGAAGACCA AGGG	21
	5569.2:	TTGGTCTTCT GGGGCTTGTT C-FL	22
	5569.3:	ACAGAGGGAG CCATACAATG AA-BIO	23
	5569.4:	BIO-TTCATTGTAT GGCTCCCTCT GTGG	24
15	1450.1:	FL-GCATGCAGGG CCTATTGCAC CAGG	25
	1450.2:	TGGTGCAATA GGCCCTGCAT G-FL	26
	1450.3:	AAATGAGAGA ACCAAGGGGA AG-BIO	27
	1450.4:	BIO-ACTTCCCCCTT GGTTCTCTCA TTTGG	28
20	1573.1:	FL-AGAAATCTAT AAAAGATGGA TAAT	29
	1573.2:	TTATCCATCT.TTTATAGATT TCT-FL	30
	1573.3:	TGGGATTAAA TAAAAAGTA AG-BIO	31
	1573.4:	BIO-CTTACTATTT TATTTAATCC CAGG	32
25	4060.1:	FL-GCATTAGGAA TCATTCAAGC ACAA	33
	4060.2:	GTGCTTGAAT GATTCCTAAT GC-FL	34
	4060.3:	AGATAAGAGT GAATCAGAGT TA-BIO	35
	4060.4:	BIO-TAACTCTGAT TCACTCTTAT CTGG	36
30	2683.1:	FL-AAGGAAGGGA AAATTTCAAA AATT	37
	2683.2:	TTTTTGAAAT TTTCCTTCC TT-FL	38
	2683.3:	CCTGAAAATC CATACAATAC T-BIO	39
	2683.4:	BIO-AGTATTGTAT GGATTTTCAG GCCC	40
35	1753.1:	FL-AACCTTGTTG GTCCAAAATG CAAA	41
	1753.2:	GCATTTTGGA CCAACAAGGT.T-FL	42
	1753.3:	AGATTGTAAG ACTATTTTAA A-BIO	43
	1753.4:	BIO-TTTAAAATAG TCTTACAATC TGGG	44
40	3697.1:	FL-GTAATATGGG GAAAGACTCC TAAA	45
	3697.2:	AGGAGTCTTT CCCCATATTA C-FL	46
	3697.3:	AAACTACCCA TACAAAAGGA A-BIO	47
	3697.4:	BIO-TTCCTTTTGT ATGGGTAGTT TAAA	48
45			

**Example 1.**

Double gap LCR (DG2,2) was performed for 40 cycles consisting of a 65 second incubation at 85°C and a 65 second incubation at 50°C. Reactions were set up with either 0 or 10<sup>2</sup> target molecules. The target DNA was a BamH1 fragment of HIV1 DNA cloned into pBR322. Negative reactions contained 10 nanograms of human placental DNA. Positive reactions contained the indicated number of copies of HIV1 plasmid DNA in a background of 10 nanograms of placental DNA. The LCR oligonucleotides used are listed in Table IV. These oligonucleotides are specific for the map positions 1667-

-25-

1716 within the *gag* region of the HIV1 genome (HIV SF2CG isolate, see Sanchez-Pescado). Reactions were run in a buffer containing 50mM EPPS pH 7.8, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 10 mM NH<sub>4</sub>Cl, 100 μM NAD, 10 μg/ml BSA, 5 x 10<sup>11</sup> molecules each oligonucleotides 1667.1, 1667.2, 1667.3, 1667.4 (ID Nos. 1-4, respectively), 1 μM 2'-deoxyguanosine 5'-triphosphate, 0.5 units *Thermus* DNA polymerase (Molecular Biology Resources, Inc., "MBR"), and 3400 units *Thermus thermophilus* DNA ligase. Reaction volume was 50 μl and each reaction was overlaid with 25 μl of mineral oil prior to cycling.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	5.76
15	10 <sup>2</sup>	1686.8

**Example 2.**

Double gap LCR (DG2,2) was performed for 40 cycles using the same cycling parameters described in Example 1. Reactions were set up in duplicate with either 0, 10 or 10<sup>2</sup> target DNA molecules. Negative reactions contained 1 microgram of human placental DNA. Positive reactions contained 10<sup>2</sup> HIV target DNA molecules in 1 microgram of human placental DNA. The oligonucleotides used are listed in Table IV. These oligonucleotides are specific for map positions 912-961 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reactions were run in a buffer containing 50 mM EPPS pH7.8, 20 mM K<sup>+</sup>, 30 mM MgCl<sub>2</sub>, 7.5 x 10<sup>11</sup> molecules each oligonucleotides 912.1, 912.2, 912.3, 912.4 (ID Nos. 5-8, respectively), 100 μM NAD, 1 μM 2'-deoxycytidine 5'-triphosphate, 0.5 units of *Thermus* DNA polymerase (MBR), and 3400 units *Thermus thermophilus* DNA ligase. Reaction volume was 50 μl and each reaction was overlaid with 25 μl of mineral oil.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer and amplification products were detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
35	0	3.8
	10	32.8
	10 <sup>2</sup>	604.4

-26-

**Example 3.**

Double gap LCR (DG2,2) was performed for 40 cycles consisting of a 65 second incubation at 85°C and a 65 second incubation at 60°C. Reactions were set up in quadruplicate with either 0 or 10<sup>2</sup> HIV 1 target DNA molecules. Negative reactions contained 1 microgram of human placental DNA. Positive reactions contained 10<sup>2</sup> HIV 1 target DNA molecules in 1 microgram of human placental DNA. The oligonucleotides used are listed in Table IV. These oligonucleotides are specific for map positions 2086-2135 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate) except that the nucleotide in position 15 of oligonucleotide 2086.3 was changed from G to C to disrupt a possible intramolecular hairpin structure (position 8 in oligonucleotide 2086.4 changed from C to G). Reactions were run in a buffer identical to that used in example 2 except that oligonucleotides 2086.1, 2086.2, 2086.3, and 2086.4 (ID Nos. 9-12, respectively) were used at 7.5 x 10<sup>11</sup> molecules each per reaction.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	5.9
20	10 <sup>2</sup>	249.5

**Example 4.**

Double gap LCR (DG2,2) was performed for 40 cycles using cycling parameters identical to those used in Example 3. Reactions were set up quadruplicate with either 0 or 10<sup>2</sup> HIV 1 target DNA molecules. Negative reactions contained 1 microgram of human placental DNA. Positive reactions contained 10<sup>2</sup> HIV 1 target DNA molecules in 1 microgram of human placental DNA. The oligonucleotides used are listed in Table IV. These oligonucleotides are specific for map position 789-838 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions were identical to those in Example 2 except that oligonucleotides 789.1, 789.2, 789.3, and 789.4 at 7.5 x 10<sup>11</sup> each (ID Nos. 13-16, respectively), and 2-deoxyadenosine 5'-triphosphate at 1 µM were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	3.8
	10 <sup>2</sup>	604.4

**Example 5.**

Double gap LCR (DG1,1) was performed for 45 cycles consisting of a 65 second incubation at 85°C and a 65 second incubation at 65°C. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> target DNA molecules. Negative reactions contained 10 nanograms of human placental DNA. Positive reactions contained 10<sup>2</sup> HIV 1 target DNA molecules in 10 nanograms of human placental DNA. The oligonucleotides used are specific for map positions 508-559 within the LTR region of the HIV 1 genome (HIV SF2CG isolate). Reactions were run in a buffer identical to that used in Example 1 except that oligonucleotides 508.1, 508.2, 508.3, and 508.4 (ID Nos. 17-20, respectively) were used at 5 x 10<sup>11</sup> molecules each and 2'-deoxycytidine 5'-triphosphate was used at 1 μM.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

<u>Number of target molecules</u>	<u>Rate</u>
0	4.5
10 <sup>2</sup>	1284.8

**Example 6.**

Double gap LCR (DG2,3) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table IV, and are specific for map positions 5569-5616 within the *vpu* coding region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions are like those described in Examples 1 - 3, except that approximately 5 x 10<sup>11</sup> molecules each of oligonucleotides 5569.1, 5569.2, 5569.3, 5569.4 (ID Nos. 21-24, respectively, and approximately 1 μM 2'-deoxycytidine 5'-triphosphate are used.

Following amplification, the reaction is diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

**Example 7.**

Double gap LCR (DG2,2) was performed for 45 cycles consisting of a 65 second incubation at 90°C and a 65 second incubation at 65°C. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> HIV 1 target DNA molecules. Negative reactions contained 10 nanograms of human placental DNA. Positive reactions contained 10<sup>2</sup> HIV 1 target molecules in 10 nanograms of placental DNA. The oligonucleotides used

-28-

are listed in Table IV. These oligonucleotides are specific for map positions 1450-1498 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reactions were run in a buffer identical to that described in Example 2 except that  $5 \times 10^{11}$  molecules each of oligonucleotides 1450.1, 1450.2, 1450.3, and 1450.4 (ID Nos. 25-28, respectively) were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

10	<u>Number of target molecules</u>	<u>Rate</u>
	0	4.8
	102	1932.5

#### Example 8.

Double gap LCR (DG2,1: 2 dNTPs) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table IV, and are specific for map positions 1573-1620 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions are similar to those described in Examples 1 - 3, except that approximately  $5 \times 10^{11}$  molecules each of oligonucleotides 1573.1, 1573.2, 1573.3, 1573.4 (ID Nos. 29-31, respectively), and approximately 1  $\mu$ M each 2'-deoxycytidine 5'-triphosphate and 2'-deoxyadenosine 5'-triphosphate are used.

Following amplification, reaction is diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

#### Example 9.

Double gap LCR (DG2,2: 2 dNTPs) are performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table IV, and are specific for map positions 4060-4107 within the *pol* region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions are similar to those described in Examples 1 - 3, except that approximately  $5 \times 10^{11}$  molecules each of oligonucleotides 4060.1, 4060.2, 4060.3, 4060.4, (ID Nos. 33-36, respectively) and approximately 1  $\mu$ M each 2'-deoxythymidine 5'-triphosphate and 2'-deoxycytidine 5'-triphosphate are used.

Following amplification, reaction are diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

-29-

**Example 10.**

Double gap LCR (DG3,2:2 dNTPs) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table IV, and are specific for map positions 2683-2730 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions are similar to those described in Examples 1 - 3, except that approximately  $5 \times 10^{11}$  molecules each of oligonucleotides 2683.1, 2683.2, 2683.3, 2683.4 (ID Nos. 37-40, respectively), and approximately 1  $\mu$ M each 2'-deoxyadenosine 5'-triphosphate and 2'-deoxyguanosine 5'-triphosphate are used.

Following amplification, reaction is diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

**Example 11.**

Double gap LCR (DG3,3:2dNTPs) is performed for 30-50 cycles using the oligonucleotides listed in Table IV. These oligonucleotides are specific for map positions 1753-1800 within the *gag* region of the HIV 1 genome (HIV SF2CG isolate). Reaction conditions are similar to those described in Examples 1 -3, except that approximately  $5 \times 10^{11}$  molecules each of oligonucleotides 1753.1, 1753.2, 1753.3, 1753.4 (ID Nos. 41-44, respectively), and approximately 1  $\mu$ M 2'-deoxythymidine 5'-triphosphate and 2'-deoxycytidine 5'-triphosphate are used.

Following amplification, reaction is diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

**Example 12.**

Double gap LCR (DG3,3) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table IV, and are specific for map positions 3697-3744 within the *pol* region of the HIV 1 genome (HIV1 SF2CG isolate) except that the nucleotide in position 17 of oligonucleotide 3697.1 was changed from T to C to disrupt a possible intramolecular hairpin structure (position 5 in oligonucleotide 3697.2 changed from A to G). Reaction conditions are like those described in Examples 1-3, except that approximately  $5 \times 10^{11}$  molecules each oligonucleotide 3697.1, 3697.2, 3697.3, 3697.4 (ID Nos. 45-48, respectively), and 1  $\mu$ M 2'-deoxythymidine 5'-triphosphate are used.

Following amplification, the reactions are diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

-30-

Examples 13-15 relate to the detection of HIV 2 DNA. The sequences, their map position according to Franchini, G., et al. *Proc. Natl. Acad. Sci.* 86:2433-2437 (1988), and their Sequence ID Nos. are given in Table V, below.

5	TABLE V		
	MAP POS	SEQUENCE	SEQ ID No.
10	2060.1:	FL-GACAGAGGAC TTGCTGCACC TCAA	49
	2060.2:	GAGGTGCAGC AAGTCCTCTG TCG-FL	50
	2060.3:	CTCTCTTTGG AAAAGACCAG TAG-BIO	51
	2060.4:	BIO-TACTGGTCTT TTCAAAGAG AGAA	52
15	4661.1:	FL-ATCACCACAG AACAAGAAAT ACAA	53
	4661.2:	GTATTTCTTG TTCTGTGGTG ATC-FL	54
	4661.3:	CCTCCAAGCA AAAAATTCAA AAT-BIO	55
	4661.4:	BIO-TTTTGAATTT TTTGCTTGA GGAA	56
20	1094.1:	FL-CTATGATATT AATCAAATGC TTAA	57
	1094.2:	AAGCATTTGA TTAATATCAT AGG-FL	58
	1094.3:	GTGTGGGCGA CCATCAAGCA GCG-BIO	59
	1094.4:	BIO-GCTGCTTGAT GGTCGCCCAC ACAA	60

### Example 13.

25 Double gap LCR (DG2,2) was performed for 45 cycles using cycling parameters identical to those described in Example 1. Reactions were set up in triplicate with either 0,  $10^2$ , or  $10^3$  target molecules. The target DNA was a segment of HIV 2 genomic DNA cloned into pTZ18. Negative reactions contained 1 microgram of human placental DNA. Positive reactions contained the indicated number of copies of HIV 2

30 containing plasmid DNA in a background of 1 microgram of human placental DNA. The oligonucleotides used are listed in Table V. These oligonucleotides are specific for map positions 2060-2109 within the *gag* region of the HIV 2 genome (HIV21SY isolate) except that the nucleotide in position 14 of oligonucleotide 2060.3 was changed from C to A to disrupt a possible intramolecular hairpin structure (position 9 in

35 oligonucleotide 2060.4 changed from G to T).. Reactions were run in a buffer identical to that described in Example 2 except that  $5 \times 10^{11}$  molecules each oligonucleotides 2060.1, 2060.2, 2060.3, 2060.4 (ID Nos.49-52, respectively), and 1  $\mu$ M 2'-deoxythymidine 5'-triphosphate were used.

40 Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.



-31-

<u>Number of target molecules</u>	<u>Rate</u>
0	8.4
10 <sup>2</sup>	318.4
10 <sup>3</sup>	789.3

5

**Example 14.**

Double gap LCR (DG2,2) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table V, and are specific for map positions 1094-1143 within the *gag* region of the HIV 2 genome (HIV21SY isolate).

10 Reaction conditions are like those described in Examples 1-3, except that approximately  $5 \times 10^{11}$  molecules each oligonucleotides 1094.1, 1094.2, 1094.3, 1094.4 (ID Nos. 53-56, respectively), and 1  $\mu$ M 2'-deoxythymidine 5'-triphosphate are used.

15 Following amplification, the reactions are diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

**Example 15.**

20 Double gap LCR (DG2,2) is performed for 30-50 cycles as in the previous examples. The oligonucleotides used are listed in Table V, and are specific for map positions 4661-4710 within the HIV 2 genome (HIV21SY isolate). Reaction conditions are like those described in Examples 1-3, except that approximately  $5 \times 10^{11}$  molecules each oligonucleotides 4661.1, 4661.2, 4661.3, 4661.4 (ID Nos. 57-60, respectively), and 1  $\mu$ M 2'-deoxythymidine 5'-triphosphate are used.

25 Following amplification, the reactions are diluted 1:1 with IMx diluent buffer and amplification products are detected via a sandwich immunoassay performed with the Abbott IMx automated immunoassay system.

30 Examples 16 - 19 relate to detection of Human Papilloma Virus (HPV), types 16 and 18. The sequences and their Sequence ID Nos are given in Table VI. Map positions for type 18 (Examples 8 and 10) are found in Seedorf, K., et al. *Virology* 145:181-185 (1985); while map positions for type 16 are found in Zhang, Y.-X., Morrison, S.G., and H.D. Caldwell, *Nuc. Acids Res.* 18:1061 (1990).

-32-

Table VI

	MAP POS	SEQUENCE	SEQ ID No.
5	631.1:	FL-TTTAGAGCCC CAAAATGAAA TTCC	61
	631.2:	AATTTCATTT TGGGGCTCTA AA-FL	62
	631.3:	TTGACCTTCT ATGTCACGAG -BIO	63
	631.4:	BIO-CTCGTGACAT AGAAGGTCAA CC	64
10	480.1:	FL-ATAATATAAG GGGTCGGTGG ACC	65
	480.2:	TCCACCGACC CCTTATATTA T-FL	66
	480.3:	TCGATGTATG TCTTGTTGCA GA-BIO	67
	480.4:	BIO-TCTGCAACAA GACATACATC GACC	68
15	488.1:	FL-CAACATAGCT GGGCACTATA GAGG	69
	488.2:	TCTATAGTGC CCAGCTATGT TG-FL	70
	488.3:	AGTGCCATTG GTGCTGCAAC C-BIO	71
	488.4:	BIO-GGTTGCAGCA CGAATGGCAC TGG	72
20	6604.1	FL-TTTGTTGGGG TAATCAATTA TTTGTT	73
	6604.2	CAAATAATTG ATTACCCCAA CAAA-FL	74
	6604.3	CTGTGGTTGA TACCACACGC A-BIO	75
	6604.4	BIO-TGCGTGTGGT ATCAACCACA GT	76

#### 25 Example 16.

Double gap LCR was performed for 35 cycles using the same cycling parameters described in Example 1. Reactions were set up in with either 0 or  $10^2$  target DNA molecules. The target DNA was full length HPV 18 genomic DNA cloned into pGEM3. Negative reactions contained 15 nanograms of calf thymus DNA. Positive reactions contained  $10^2$  HPV 18 target DNA molecules in 15 nanograms of calf thymus DNA. The oligonucleotides used are listed in Table VI. These oligonucleotides are specific for map positions 631-676 in the E6/E7 region of the HPV 18 genome<sup>3</sup>. Reactions were run in a buffer identical to that described in Example 1 except that  $5 \times 10^{11}$  molecules each oligonucleotides 631.1, 631.2, 631.3, 631.4 (ID Nos. 61-64, respectively) and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	6.8
	$10^2$	166.6

-33-

**Example 17.**

Double gap LCR (DG2,2) was performed for 35 cycles using reaction conditions identical to those described in Example 3. Reactions were set up in with either 0 or  $10^2$  target DNA molecules. The target DNA was full length HPV 16 genomic DNA cloned into pSP65. Negative reactions contained 15 nanograms of calf thymus DNA. Positive reactions contained  $10^2$  target HPV 16 DNA molecules in 15 nanograms of calf thymus DNA. The oligonucleotides used are listed in Table VI. These oligonucleotides are specific for map positions 480-526 within the E6/E7 region of the HPV 16 genome. Reactions were run in a buffer identical to that used in Example 1 except that  $5 \times 10^{11}$  molecules each oligonucleotides 480.1, 480.2, 480.3, 480.4 (ID Nos. 65-68, respectively), and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

<u>Number of target molecules</u>	<u>Rate</u>
0	5.9
$10^2$	289.6

**Example 18.**

Double gap LCR (DG2,2) was performed for 35 cycles using the same cycling parameters described in Example 1. Reactions were set up in duplicate with either 0 or  $10^6$  target DNA molecules. The target DNA was full length HPV 18 genomic DNA cloned into a plasmid. Negative reactions contained 15 nanograms of calf thymus DNA. Positive reactions contained  $10^6$  target HPV 18 DNA molecules in 15 nanograms of calf thymus DNA. The oligonucleotides used are listed in Table VI. These oligonucleotides are specific for map positions 488-534 within the E6/E7 region of the HPV 18 genome. Reactions were run in a buffer identical to that described in Example 1 except that  $5 \times 10^{11}$  molecules each oligonucleotides 488.1, 488.2, 488.3, 488.4 (ID Nos. 69-72, respectively), and  $1 \mu\text{M}$  2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

-34-

<u>Number of target molecules</u>	<u>Rate</u>
0	7.0
10 <sup>6</sup>	978.2

5 Example 19.

Double gap LCR (DG1,2) was performed for 40 cycles consisting of a 65 second incubation at 88°C and an 80 second incubation at 50°C. Reactions were set up in triplicate with either 0, 10<sup>4</sup>, or 10<sup>6</sup> target molecules. The target DNA was full length viral genomes cloned into various plasmids. All reactions contained 330 nanograms of human placental DNA. Positive reactions contained 10<sup>4</sup> or 10<sup>6</sup> molecules of HPV 6, 16, 18, 31, 33, or 61. The oligonucleotides used are listed in Table VI. These oligonucleotides correspond to map positions 6604-6651 in the HPV 16 genome, but do not match this sequence exactly. Nucleotide changes representing consensus between all of the HPV types listed above were introduced during synthesis. Reactions were run in a buffer identical to that described in Example 2 except that 4 x 10<sup>12</sup> molecules each of oligonucleotides 6604.1, 6604.2, 6604.3, 6604.4 (ID Nos. 73-76 respectively), 18,000 units of *Thermus thermophilus* DNA ligase, 4 units of *Thermus* DNA polymerase (MBR), 10μM NAD, and 1.7μM 2'-deoxyadenosine 5'-triphosphate were used. Reaction volume was 200 μl with no mineral oil overlay.

Following amplification, LCR reaction products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

Rate (c/s/s)						
<u>No. of target molecules</u>	<u>HPV Type 6</u>	<u>HPV Type 16</u>	<u>HPV Type 18</u>	<u>HPV Type 31</u>	<u>HPV Type 33</u>	<u>HPV Type 61</u>
0	6.6	6.6	6.6	6.6	6.6	6.6
10 <sup>4</sup>	355.4	31.4	15.3	66.9	78.8	7.4
10 <sup>6</sup>	2173.7	1163.8	805.5	1439.7	1597.8	62.1

25 Examples 20 - 21 relate to detection of Herpes Simplex Virus (HSV). The sequences used, their map position according to McGeoch, D.J. et al. *J. Gen. Virol.* 68:19-38 (1987), and their Sequence ID Nos. are given in Table V.

-35-  
Table VII.

	MAP POS	SEQUENCE	SEQ ID No.
5	4751.1:	FL-CCCCCTGTTC TGGTTCCTAA CGG	77
	4751.2:	GTTAGGAACC AGAACAGTGG GGA-FL	78
	4751.3:	TCCCCTGCTC TAGATATCCT CT-BIO	79
	4751.4:	BIO-GAGGATATCT AGAGCAGGGG AGG	80
10	6465.1:	FL-TATGACAGCT TTAGCGCCGT CAG	81
	6465.2:	TGACGGCGCT AAAGCTGCAT AG-FL	82
	6465.3:	GAGGATAACC TGGGGTTCCT GAT-BIO	83
	6465.4:	BIO-TCAGGAACCC CAGGTTATCC TCG	84

#### 15 Example 20.

Double gap LCR (DG2,2) was performed for 60 cycles consisting of a 60 second incubation at 85°C and a 60 second incubation at 50°C. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> target DNA molecules. The target DNA was a 4538 bp segment of the HSV 2 genome cloned into pUC19. Negative reactions contained 10 nanograms of human placental DNA. Positive reactions contained 10<sup>2</sup> HSV 2 plasmid molecules in 10 nanograms of human placental DNA. The oligonucleotides used are listed in Table VII. These oligonucleotides are specific for map positions 4751-4798 within the US4 region of the HSV 2 genome<sup>2</sup>. Reactions were run in a buffer identical to that described in Example 1 except that oligonucleotides 4751.1, 4751.2, 4751.3, and 4751.4 (ID Nos. 77-80, respectively) at 5 x 10<sup>11</sup> molecules each per reaction, and 1 μM 2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

30

<u>Number of target molecules</u>	<u>Rate</u>
0	6.0
10 <sup>2</sup>	2096.0

#### 35 Example 21.

Double gap LCR (DG1,1) was performed for 60 cycles using cycling parameters identical to those described in Example 1. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> target DNA molecules. Negative reactions contained 10 nanograms of human placental DNA. Positive reactions contained 10<sup>2</sup> HSV target DNA molecules in 10 nanograms of human placental DNA. The oligonucleotides used are listed in Table VII. These oligonucleotides are specific for map positions 6465-6512 of the HSV 2 genome. Reactions were run in a buffer identical to that described in Example 1

40

-36-

except that oligonucleotides 6465.1, 6465.2, 6465.3, and 6465.4 (ID Nos. 81-84, respectively) at  $5 \times 10^{11}$  each per reaction and  $1 \mu\text{M}$  2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and  
 5 LCR reaction products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	6.0
10	$10^2$	1156.0

Examples 22 - 24 relate to detection of *Chlamydia trachomatis* DNA. The sequences used and their Sequence ID Nos. are given in Table VII. The map positions for the MOMP region (Examples 11 and 12) are according to Baehr, W. et al. *Proc. Natl. Acad. Sci. USA* 85:4000-4004 (1988) [EMBL Genbank Data Base:  
 15 Intelligenetics, Accession #J03813]; while the map positions for the cryptic plasmid region (Example 13) are according to Hatt, C. et al. *Nuc. Acids Res.* 16:4053-4067 (1988).

20 Table VIII.

	<u>MAP POS</u>	<u>SEQUENCE</u>	<u>SEQ ID No.</u>
	36.1:	FL-TTTTACTTGC AAGACATTCC TCAGG	85
25	36.2:	TGAGGAATGT CTTGCAAGTA AAAGC-FL	86
	36.3:	ATTAATTGCT ACAGGACATC TTGTC-BIO	87
	36.4:	BIO-CAAGATGTCC TGTAGCAATT AATGG	88
	552.1:	GGGAATCCTG CTGAACCAAG	89
30	552.2:	TTGGTTCAGC AGGATTCCC	90
	552.3:	TTATGATCGA CGGAATTCTG TG	91
	552.4:	CACAGAATTC CGTCGATCAT AAGG	92
	6693.1:	FL-GATACTTCGC ATCATGTGTT CC	93
35	6693.2:	AACACATGAT GCGAAGTATC -FL	94
	6693.3:	AGTTTCTTTG TCCTCCTATA ACG-BIO	95
	6693.4:	BIO-CGTTATAGGA GGACAAAGAA ACTCC	96

#### Example 22.

40 Double gap LCR (DG2,2) was performed for 35 cycles using cycling parameters identical to those described in Example 3. Reactions were set up in duplicate with either 0 or  $10^2$  target DNA molecules. The target DNA was purified *Chlamydia trachomatis* genomic DNA. Negative reactions contained 330 nanograms of human placental DNA. Positive reactions contained  $10^2$  *Chlamydia* target DNA molecules in

-37-

330 nanograms of human placental DNA. The oligonucleotides used are listed in Table VIII. These oligonucleotides are specific for map positions 36-89 within the major outer membrane protein (MOMP) gene. Reactions were run in a buffer identical to that described in Example 2 except that  $5 \times 10^{11}$  molecules each oligonucleotides

5 36.1, 36.2, 36.3, 36.4 (ID Nos. 85-88, respectively) and  $1 \mu\text{M}$  2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

10

<u>Number of target molecules</u>	<u>Rate</u>
0	7.2
$10^2$	34.2

#### 15 Example 23.

Double gap LCR (DG1,2) was performed for 40 cycles consisting of a 60 second incubation at  $85^\circ\text{C}$  and a 60 second incubation at  $59^\circ\text{C}$ . Reactions were set up in duplicate with either 0 or  $10^2$  target molecules. Negative reactions contained 330 nanograms of human placental DNA. Positive reactions contained  $10^2$  *Chlamydia*

20 target DNA molecules in 330 nanograms of human placental DNA. The oligonucleotides used are listed in Table VIII. These oligonucleotides are specific for map positions 552-595 within the MOMP gene. Reactions were run in a buffer identical to that described in Example 3 except that  $8 \times 10^{11}$  molecules each of oligonucleotides 552.1 552.2 552.3 552.4 (ID Nos. 89-92, respectively), and  $1 \mu\text{M}$   $\alpha$   $^{32}\text{P}$  labelled

25 2'-deoxycytidine 5'-triphosphate are used.

Following amplification, the reaction products were separated by electrophoresis on a 10% polyacrylamide gel with 8.0M urea at 330 v constant for about 2 hours. The gels were then autoradiographed for about 12 hours with intensifying screens (optional). In the presence of target, autoradiography revealed bands equatable with

30 the extended, unligated probe, and the longer, ligated product. The negative control gave no discernable bands.

#### Example 24.

Double gap LCR (DG2,2) was performed for 40 cycles consisting of a 60 second

35 incubation at  $85^\circ\text{C}$  and a 60 second incubation at  $59^\circ\text{C}$ . Reactions were set up in duplicate with either 0 or  $10^2$  target molecules. Negative reactions contained 330 nanograms of human placental DNA. Positive reactions contained  $10^2$  *Chlamydia* target DNA molecules in 330 nanograms of human placental DNA. The oligonucleotides

-38-

used are listed in Table VIII. These oligonucleotides are specific for map positions 6693-6739 within the *Chlamydia trachomatis* cryptic plasmid. Reactions were run in a buffer identical to that described in Example 3 except that  $8 \times 10^{11}$  molecules each oligonucleotides 6693.1, 6693.2, 6693.3, 6693.4 (ID Nos. 93-96 respectively), and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
	0	4.8
	$10^2$	1538.8

Examples 25 - 29 relate to detection of *Neisseria gonorrhoeae*. The sequences used are given in Table IX. The map positions for the *laz* region (example 25) are according to Gotschlich, E. C., and M. E. Seiff. *FEMS Microbiol. Letts.* 43: 253-255 (1987). The map positions for the P.II/*opa* region (examples 26 & 27) are according to Stern, A., M. Brown, P. Nickel, and T. F. Meyer. *Cell* 47: 61-71 (1986). The map positions for the *pilIn* region (examples 28 & 29) are according to Meyer, T. F., E. Billyard, R. Haas, S. Storzbach, and M. So. *Proc. Natl. Acad. Sci. USA* 81: 6110-6114 (1984).

TABLE IX

	<u>MAP POS</u>	<u>SEQUENCE</u>	<u>SEQ ID No.</u>	<u>GENE</u>
25	27.1:	FL-AATATTTGCG CGGCGGGGAC GG	97	<i>laz</i>
	27.2:	GTCCCCGCCG CGCAAATATT -FL	98	
	27.3:	TGCCCTAATA TTAAAGAAAT AGGTT-BIO	99	
30	27.4:	BIO-AACCTATTTC TTTAATATTA GGCAAG	100	
	66.1:	FL-GCCATATTGT GTTGAAACAC CGCCC	101	<i>opa</i>
	66.2:	CGGTGTTTCA ACACAATATG GC-FL	102	
	66.3:	AACCCGATAT AATCCGCCCT T-BIO	103	
35	66.4:	BIO-AAGGGCGGAT TATATCGGGT TCC	104	
	114.1:	FL-CAACATCAGTG AAAATCTTTT TTAAACC	105	<i>opa</i>
	114.2:	TTAAAAAAG ATTTTCACTG ATGTTG-FL	106	
	114.3:	TCAAACCGAA TAAGGAGCCG A-BIO	107	
40	114.4:	BIO-TTCGGCTCCT TATTCGGTTT GACC	108	



-39-

5	822.1:	FL-TGATGCCAGC TGAGGCAAAT TAGG	109	pilin
	822.2:	TAATTTGCCT CAGCTGGCAT CA-FL	110	
	822.3:	TTAAATTTCA AATAAATCAA GCGGTA-BIO	111	
	822.4:	BIO-TACCGCTTGA TTTATTGAA ATTTAAGG	112	
10	933.1:	FL-CGGGCGGGGT CGTCCGTTCC	113	pilin
	933.2:	AACGGACGAC CCCGCCG-FL	114	
	933.3:	TGGAAATAAT ATATCGATT-BIO	115	
	933.4:	BIO-AATCGATATA TTATTTCCAC C	116	

<sup>a</sup> The sequences shown are a consensus of all published *N. gonorrhoeae* pilin genes and when taken in total do not exactly match any one particular sequence.

15

**Example 25.**

Double gap LCR (DG2,1) was performed for 35 cycles consisting of a 30 second incubation at 85°C and a 60 second incubation at 54°C. Reactions were set up in duplicate with genomic DNA from the indicated cells. All *Neisseria* DNA's were tested in the presence of 320 ng human placenta DNA. The oligonucleotides used (see Table IX, above) are specific for map positions 27-74 within the *Neisseria gonorrhoeae* *laz* gene. Reactions were run in a buffer identical to that described in Example 3 except that  $5 \times 10^{11}$  molecules each of oligonucleotides 27.1, 27.2, 27.3, 27.4 (ID Nos. 97-100, respectively), and  $1 \mu\text{M}$  2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

30

	<u>DNA Source (Amount)</u>	<u>Rate</u>
35	<i>Neisseria gonorrhoeae</i> (200 cells)	576
	<i>Neisseria meningitidis</i> ( $2 \times 10^6$ cells)	6
	<i>Neisseria lactamica</i> ( $2 \times 10^6$ cells)	5
	Human placenta (320 ng)	5

**Example 26.**

Double gap LCR (DG3,2) was performed for 27 cycles consisting of a 30 second incubation at 85°C and a 60 second incubation at 57°C. Reactions were set up in duplicate with genomic DNA from the indicated cells. All *Neisseria* DNA's were tested in the presence of 320 ng human placenta DNA. The oligonucleotides used (see Table IX, above) are specific for map positions 66-113 within the *Neisseria gonorrhoeae* *opa* gene. Reactions were run in a buffer identical to that described in

40

-40-

Example 3 except that  $5 \times 10^{11}$  molecules each of oligonucleotides 66.1, 66.2, 66.3, 66.4 (ID Nos. 101-104, respectively), and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>DNA Source (Amount)</u>	<u>Rate</u>
10	<i>Neisseria gonorrhoeae</i> (200 cells)	1710
	<i>Neisseria meningitidis</i> ( $2 \times 10^6$ cells)	18
	<i>Neisseria lactamica</i> ( $2 \times 10^6$ cells)	8
	Human placenta (320 ng)	6

#### 15 Example 27.

Double gap LCR (DG2,2) was performed for 35 cycles consisting of a 30 second incubation at  $85^\circ\text{C}$  and a 60 second incubation at  $54^\circ\text{C}$ . Reactions were set up in duplicate with genomic DNA from the indicated cells. All *Neisseria* DNA's were tested in the presence of 320 ng human placenta DNA. The oligonucleotides used (see Table IX, above) are specific for map positions 114-165 within the *Neisseria gonorrhoeae* opa gene. Reactions were run in a buffer identical to that described in Example 3 except that  $5 \times 10^{11}$  molecules each of oligonucleotides 114.1, 114.2, 114.3, 114.4 (ID Nos. 105-108, respectively), and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>DNA Source (Amount)</u>	<u>Rate</u>
30	<i>Neisseria gonorrhoeae</i> (200 cells)	306
	<i>Neisseria meningitidis</i> ( $2 \times 10^6$ cells)	10
	<i>Neisseria lactamica</i> ( $2 \times 10^6$ cells)	10
	Human placenta (320 ng)	9

35

#### Example 28.

Double gap LCR (DG2,2) was performed for 35 cycles consisting of a 30 second incubation at  $85^\circ\text{C}$  and a 60 second incubation at  $50^\circ\text{C}$ . Reactions were set up in duplicate with genomic DNA from the indicated cells. All *Neisseria* DNA's were tested in the presence of 320 ng human placenta DNA. The oligonucleotides used (see Table IX, above) are specific for map positions 822-873 within the *Neisseria gonorrhoeae* pilin gene. Reactions were run in a buffer identical to that described in

-41-

Example 3 except that  $5 \times 10^{11}$  molecules each of oligonucleotides 822.1, 822.2, 822.3, 822.4 (ID Nos. 109-112, respectively), and  $1 \mu\text{M}$  2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>DNA Source (Amount)</u>	<u>Rate</u>
10	<i>Neisseria gonorrhoeae</i> (200 cells)	1291
	<i>Neisseria meningitidis</i> ( $2 \times 10^6$ cells)	27
	<i>Neisseria lactamica</i> ( $2 \times 10^6$ cells)	22
	Human placenta (320 ng)	11

#### 15 Example 29.

Double gap LCR (DG2,2) is performed for 30 - 50 cycles as in earlier examples. The oligonucleotides used (see Table IX, above) are specific for map positions 933-973 within the *Neisseria gonorrhoeae* pilin gene. Reactions are run in a buffer identical to that described in Example 3 except that approximately  $5 \times 10^{11}$  molecules each of oligonucleotides 933.1, 933.2, 933.3, 933.4 (ID Nos. 113-116, respectively), and  $1 \mu\text{M}$  2'-deoxyguanosine 5'-triphosphate are used.

Following amplification, reactions are diluted 1:1 with IMx diluent buffer, and the LCR amplification products are detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

25

Examples 30 and 31 relate to the detection of *Borrelia burgdorferi*, the causative agent of "Lyme" disease. The sequences used, their map position according to Rosa, P.A., and T.G. Schwan. *J. Infect. Dis.* 160:1018-1029 (1989); and their Sequence ID Nos. are given in Table X, below.

Table X.

	<u>MAP POS</u>	<u>SEQUENCE</u>	<u>SEQ ID No.</u>
35	5.1:	FL-AAAACGAAGA TACTAAATCT GTAA	117
	5.2:	ACAGATTTAG TATCTTCGTT TT-FL	118
	5.3:	GCAGAAACAC CTTTGAATT AA-BIO	119
	5.4:	BIO-TTAATTCAAA AGGTGTTTCT GCAA	120
40	181.1:	FL-CATCTTTTGG AGCTAAATAT AAG	121
	181.2:	TTATATTTAG CTCCAAAAGATGC-FL	122
	181.3:	TTGGATTAAC AAAAATAAAC GAT-BIO	123
	181.4:	BIO-TCGTTTATTT TTGTTAATCC AAG	124

Example 30.

Double gap LCR (DG2,2) was performed for 45 cycles consisting of a 25 second incubation at 85°C and a 70 second incubation at 50°C. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> target molecules. The target DNA was a random *Borrelia burgdorferi* genomic sequence cloned into pUC18. Negative reactions contained 10 nanograms of calf thymus DNA. Positive reactions contained 10<sup>2</sup> *Borrelia* target DNA molecules in 10 nanograms of calf thymus DNA. The oligonucleotides used are listed in Table X, and are specific for map positions 5-53 of the cloned *Borrelia* genomic DNA. Reactions were run in a buffer identical to that described in Example 1 except that 3 x 10<sup>12</sup> molecules each oligonucleotides 5.1 and 5.3 (ID Nos. 117 and 119, respectively), 2 x 10<sup>12</sup> molecules each oligonucleotides 5.2 and 5.4 (ID Nos. 118 and 120, respectively), 1032 units of *Thermus thermophilus* DNA ligase, and 1 μM 2'-deoxythymidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

	<u>Number of target molecules</u>	<u>Rate</u>
20	0	6.0
	10 <sup>2</sup>	152.0

Example 31.

Double gap LCR (DG1,1) was performed for 45 cycles consisting of a 65 second incubation at 85°C and a 60 second incubation at 50°C. Reactions were set up in duplicate with either 0 or 10<sup>2</sup> target molecules. Negative reactions contained 10 nanograms of salmon sperm DNA. Positive reactions contained 10<sup>2</sup> *Borrelia* target DNA molecules in 10 nanograms of salmon sperm DNA. The oligonucleotides used are listed in Table X, and are specific for map positions 181-228 of the cloned *Borrelia* genomic DNA. Reactions were run in a buffer identical to that described in Example 1 except that 1 x 10<sup>12</sup> molecules each oligonucleotides 181.1, 181.2, 181.3, 181.4 (ID Nos. 121-124, respectively), 500 units of *Thermus thermophilus* DNA ligase, 1.6 units of *Thermus* DNA polymerase (MBR), and 1 μM 2'-deoxycytidine 5'-triphosphate were used.

Following amplification, reactions were diluted 1:1 with IMx diluent buffer, and the LCR amplification products were detected via a sandwich immunoassay performed using the Abbott IMx automated immunoassay system.

-43-

Number of target moleculesRate

0

5.0

10<sup>2</sup>

319.0

- 5        The above examples serve to illustrate the invention, but are not intended to limit it in any way, the full scope of the invention being defined by the appended claims.

-44-

What is claimed is:

1. A method for detecting the presence of a target nucleic acid sequence in a sample, said method employing the ligase chain reaction to create geometrically increasing numbers of reorganized probe molecules in the presence of said target sequence, said method comprising:
  - 5 a) providing a sample suspected to contain nucleic acid, the nucleic acid having a target sequence of the formula:
 
$$5'-(N)_h X E_p F_q Z (N)_k -3'$$
 wherein E represents any base, F represents any base except E, p and q are independently integers from 1 to 10, X represents any base except E or F', Z represents any base except F or E', N represents any base, and h and k are independently integers from 5 to 100, provided the sums (p+h) and (q+k) each are greater than 10, wherein (N)<sub>h</sub> and (N)<sub>k</sub> represent sequences that are characteristic for the target sought to be detected;
    - 15 b) providing a plurality of each of four probes having the formulas:
 
$$\begin{aligned} A: & 5'-(N)_{h'} X E_p -3' \\ A': & 3'-(N')_{h'} X' -5' \\ B: & 5'-Z (N)_{k'} -3' \\ B': & 3'-F'_q Z' (N')_{k'} -5' \end{aligned}$$
 wherein the symbols have the same meanings given above, except that h' and k' need only approximate h and k, respectively, varying by no more than 25%, and wherein at least one of the probes is labeled with a moiety capable of detection; and also providing deoxyribonucleotide triphosphates of E' and F, a polymerase reagent and a ligase reagent;
      - 25 c) performing the following cycle at least once:
        - i) mixing said probes with said sample under hybridizing conditions to allow probes to hybridize to the target sequence and its complement if present, or to reorganized probes created therefrom;
        - ii) using target sequence or reorganized probes created therefrom as template, extending probe A with said polymerase reagent by adding F deoxyribonucleotide triphosphates to its 3' end, and extending probe B' with said polymerase reagent by adding E' deoxyribonucleotide triphosphates to its 3' end
        - 30 iii) ligating extended probe A to probe B, and extended probe B' to probe A', using said ligase reagent to form reorganized probe molecules; and
        - 35 iv) providing denaturing conditions to separate said reorganized probe molecules from said template;
      - d) separating reorganized probe molecules from unreorganized labeled probes; and

-45-

- e) detecting the presence of said label in the reorganized or fraction as a measure of the presence of the target sequence.
2. The method according to claim 1 wherein F is E', whereby step b) requires adding the deoxyribonucleotide triphosphate of only E'.
  - 5 3. The method according to claim 1 or 2 wherein p and q are independently integers between 1 and 3, inclusive.
  4. The method according to claim 1 wherein the 5' end of probe A and the 3' end of probe A' are labeled with a first hapten.
  5. The method according to claim 4 wherein the separation step (step d)
  - 10 comprises capturing on a solid phase said first hapten and its associated probe.
  6. The method according to claim 1 wherein the cycle of step c) is repeated from about 20 to about 60 times.

7. A diagnostic kit for the detection of a target nucleic acid sequence in a sample,
- 15 said nucleic acid having a target sequence comprising



- wherein E and F independently represent any base, p and q are independently integers from 1 to about 10, X represents any base except E or F', Z represents any base except F or E', N represents any base, and h and k are independently integers from 5 to 100, provided the sums (p+h) and (q+k) each are greater than 10, wherein (N)<sub>h</sub> and (N)<sub>k</sub> represent sequences that are characteristic for the target sought to be
- 20 detected, said kit comprising in combination:

(a) four probes, having the formulas:

- A:  $5'-(N)_h X E_p -3'$
- 25 A':  $3'-(N')_h X' -5'$
- B:  $5'-Z (N)_k -3'$
- B':  $3'-F'_q Z' (N')_k -5'$

- wherein the symbols have the same meanings given above, except that h' and k' need only approximate h and k, respectively, varying by no more than 25%, and wherein
- 30 at least one of the probes is labeled with a moiety capable of detection;

- (b) a polymerase reagent capable of extending probe A with said polymerase reagent in a target dependent manner by adding F deoxyribonucleotide triphosphates to its 3' end, and extending probe B' with said polymerase reagent in a target dependent manner by adding E' deoxyribonucleotide triphosphates to its 3' end, thereby to
- 35 render the primary probes ligatable to one another when hybridized to target and, optionally, to render the secondary probes ligatable to one another when hybridized to target;

(c) deoxyribonucleotide triphosphates of E' and F; and

-46-

(d) a ligase reagent for ligating the extended probe A to probe B, and extended probe B' to probe A', thereby to form reorganized probe molecules.

8. The kit according to claim 7, and also comprising means for separating reorganized probe molecules from unreorganized labeled probes.

5

9. A composition of matter useful for detecting the DNA of HIV 1, said composition comprising a mixture of four DNA probes, said probes being selected from the probe sets listed in Table IV and identified by sequence ID Nos: 1, 2, 3 and 4; 5, 6, 7 and 8; 9, 10, 11 and 12; 13, 14, 15 and 16; 17, 18, 19 and 20; 21, 22, 23 and 24; 10 25, 26, 27 and 28; 29, 30, 31 and 32; 33, 34, 35 and 36; 37, 38, 39 and 40; 41, 42, 43 and 44; or 45, 46, 47 and 48.

10. A composition of matter useful for detecting the DNA of HIV 2, said composition comprising a mixture of four DNA probes, said probes being selected from the probe sets listed in Table V and identified by sequence ID Nos: 49, 50, 51 and 52; 53, 54, 15 55 and 56; or 57, 58, 59 and 60.

11. A composition of matter useful for detecting the DNA of HPV, said composition comprising a mixture of four DNA probes, said probes being selected from the probe sets listed in Table VI and identified by sequence ID Nos: 61, 62, 63 and 64; 65, 66, 67 and 68; 69, 70, 71 and 72; or 73, 74, 75 and 76.

20 12. A composition of matter useful for detecting the DNA of HSV, said composition comprising a mixture of four DNA probes, said probes being selected from the probe sets listed in Table VII and identified by sequence ID Nos: 77, 78, 79 and 80; or 81, 82, 83 and 84.

13 A composition of matter useful for detecting the DNA of *Chlamydia trachomatis*, said composition comprising a mixture of four DNA probes, said probes 25 being selected from the probe sets listed in Table VIII and identified by sequence ID Nos: 85, 86, 87 and 88; 89, 90, 91 and 92; or 93, 94, 95 and 96.

14. A composition of matter useful for detecting the DNA of *Neisseria gonorrhoeae*, said composition comprising a mixture of four DNA probes, said probes being selected 30 from the probe sets listed in Table IX and identified by sequence ID Nos: 97, 98, 99 and 100; 101, 102, 103 and 104; 105, 106, 107 and 108; 109, 110, 111 and 112; or 113, 114, 115 and 116.

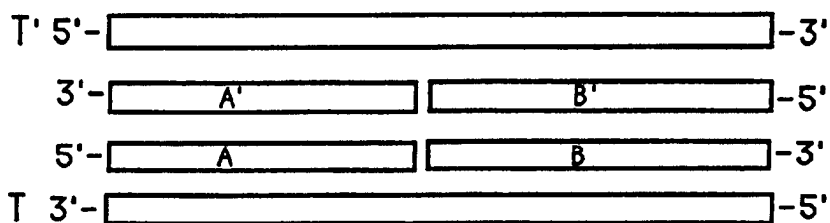
15. A composition of matter useful for detecting the DNA of *Borrelia burgdorferi*, said composition comprising a mixture of four DNA probes, said probes being selected 35 from the probe sets listed in Table X and identified by sequence ID Nos: 117, 118, 119 and 120; or 121, 122, 123 and 124.



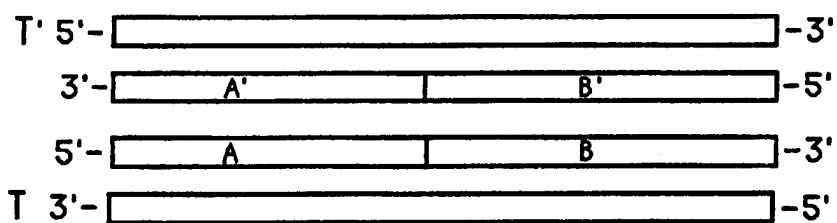
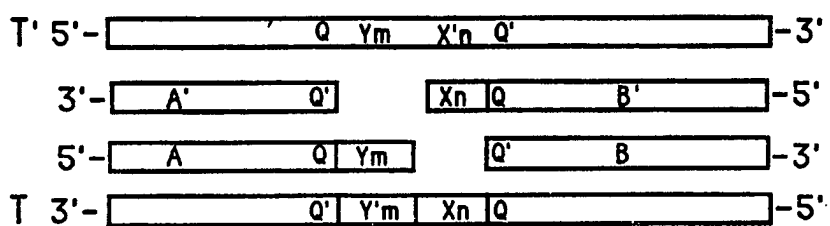
HIVSF2CG Subsequence Sites  
Thursday, June 27, 1991 8:40 AM

Sequence Range: 791 to 2299

Pos. Subseq	Start #Res.	Loc. Matched	Subseq Found	
Search with DG 2,2 (A): 6 sites found.				
811 DG 2,2 (A) (part #1)	811	6/6	+	ATTAAAG
816 DG 2,2 (A) (part #1)	816	6/6	-	CTTAAT
850 DG 2,2 (A) (part #1)	850	6/6	+	GTTAAG
855 DG 2,2 (A) (part #1)	855	6/6	-	CTTAAC
1964 DG 2,2 (A) (part #1)	1964	6/6	+	GTTAAG
1969 DG 2,2 (A) (part #1)	1969	6/6	-	CTTAAC
Search with DG 2,2 (C): 14 sites found.				
854 DG 2,2 (C) (part #1)	854	6/6	+	AGGCCA
859 DG 2,2 (C) (part #1)	859	6/6	-	TGGCCT
934 DG 2,2 (C) (part #1)	934	6/6	+	TGGCCT
939 DG 2,2 (C) (part #1)	939	6/6	-	AGGCCA
1230 DG 2,2 (C) (part #1)	1230	6/6	+	AGGCCA
1235 DG 2,2 (C) (part #1)	1235	6/6	-	TGGCCT
1471 DG 2,2 (C) (part #1)	1471	6/6	+	AGGCCA
1476 DG 2,2 (C) (part #1)	1476	6/6	-	TGGCCT
1864 DG 2,2 (C) (part #1)	1864	6/6	+	CGGCCA
1869 DG 2,2 (C) (part #1)	1869	6/6	-	TGGCCG
2108 DG 2,2 (C) (part #1)	2108	6/6	+	TGGCCT
2113 DG 2,2 (C) (part #1)	2113	6/6	-	AGGCCA
2126 DG 2,2 (C) (part #1)	2126	6/6	+	AGGCCA
2131 DG 2,2 (C) (part #1)	2131	6/6	-	TGGCCT
Search with DG 2,2 (G): 2 sites found.				
1689 DG 2,2 (G) (part #1)	1689	6/6	+	ACCGGT
1694 DG 2,2 (G) (part #1)	1694	6/6	-	ACCGGT
Search with DG 2,2 (T): 4 sites found.				
824 DG 2,2 (T) (part #1)	824	6/6	+	GAATTA
829 DG 2,2 (T) (part #1)	829	6/6	-	TAATTC
1975 DG 2,2 (T) (part #1)	1975	6/6	+	CAATTC
1980 DG 2,2 (T) (part #1)	1980	6/6	-	CAATTC

$\frac{1}{2}$ 

+ LIGASE

FIG. 1  
(Prior Art)

+ POLYMERASE  
+ LIGASE  
+ dX'TPs and dY'TPs

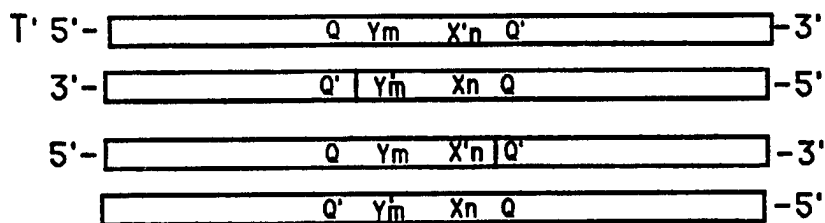
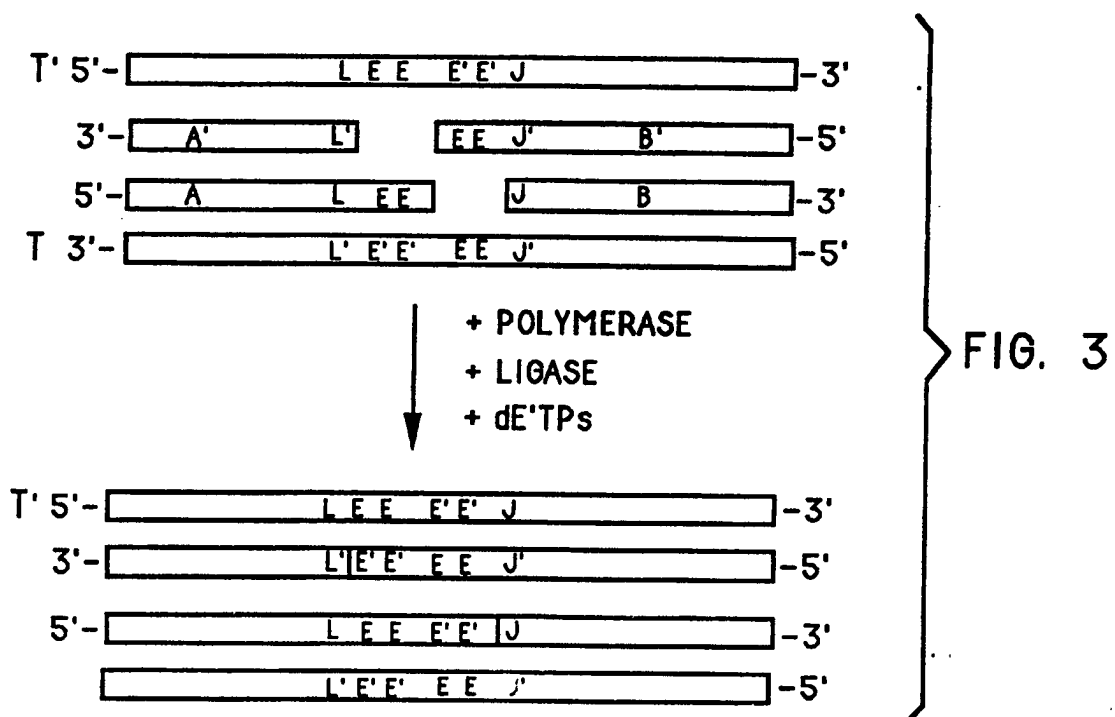


FIG. 2

2/2



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US92/05477

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : C12Q 1/68; C12P 19/34; C07H 15/12; C07H 17/00  
US CL : 435/6; 435/91; 536/27; 436/811

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 435/91; 536/27; 436/811

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAS, BIOSYS: Ligase Chain Reaction  
Ligase? w/ Chain Reaction

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP, A 0,185,494 (Whiteley et al) 25 June 1986.	1-15
Y	WO, A, 89/09835 (Orgel) 19 October 1989, See abstract.	1-15
Y	EP, A, 0,246,864 (Carr) 25 November 1987, See abstract, claims 2-4.	1-15

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
* A* document defining the general state of the art which is not considered to be part of particular relevance	* X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* E* earlier document published on or after the international filing date	* Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
* L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* A*	document member of the same patent family
* O* document referring to an oral disclosure, use, exhibition or other means		
* P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

16 SEPTEMBER 1992

Date of mailing of the international search report

07 OCT 1992

Name and mailing address of the ISA/  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

SCOTT CHAMBERS

Facsimile No. NOT APPLICABLE

Telephone No. (703) 309-0196

**BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING**

This ISA found multiple inventions as follows:

Group 1: Claims 1-9 drawn to a method of detection, classified in Class 435, subclass 6.

Group 2: Claim 10 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 1 in sequence detected, primers used, and disease condition from these of Group 1 and 3-7.

Group 3: Claim 11 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 2 in sequence detected, primers used, and disease condition from these of Groups 1-2 and 4-7.

Group 4: Claim 12 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 3 in sequence detected, primers used, and disease condition from these of Groups 1-3 and 5-7.

Group 5: Claim 13 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 4 in sequence detected, primers used, and disease condition from these of Groups 1-4 and 6-7.

Group 6: Claim 14 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 5 in sequence detected, primers used, and disease condition from these of Groups 1-5 and 7.

Group 7: Claim 15 drawn to a composition of matter, classified in 536, subclass 27 which differs from Group 6 in sequence detected, primers used, and disease condition from these of Groups 1 and 6.